MATHEMATICAL RESEARCH, INC.

HOUSTON, TEXAS 77004

NASA CR 108498

THE METHOD OF MOMENTS

THIS REPORT WAS PREPARED WITH PARTIAL SUPPORT OF NASA-MSC(HOUSTON)
CONTRACT NAS 9-6964

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

The purpose of this report is to present a detailed analysis of a special case of the Galerkin method known as the method of moments. The ideas herein are not new but represent a compilation of methods and techniques for which proofs and explaination are given in detail. The report is essentially self-contained and should serve not only as a useful introduction to one area of approximation theory but also as a guide to effective use of the method in the solution of applied problems.

Chapter II contains definitions and results which are assumed. Practically all of the theorems in the remaining Chapters are proved in detail, not only for completeness but also since many of the algorithms hinge on the ideas in the proofs.

The method of moments is described in general in Chapter III. It is also shown there how this method can be used to solve the eigenvalue problem and problems of the first and second kind in a finite dimensional space.

When the operator in question is completely continuous certain stronger results can be obtained and these are given in Chapter IV. These relate to problem solutions as well as convergence theorems. In addition it is shown how this method can be used to speed convergence in the classical Liouville-Neumann method of successive approximations.

In Chapter V results are given for self-adjoint operators in which case certain of the algorithms can be simplified.

Various types of problems frequently encountered are discussed in Chapter VI. Certain time dependent problems and problems with unbounded operators are considered.

Throughout the report following each algorthim a summary is given pointing out the necessary steps to apply the results.

CHAPTER II

PRELIMINARIES

In this section terms will be defined, special notation will be developed and theorems that will be assumed will be stated.

Definition 1 A set R of functions forms a linear space over the field of complex numbers, C, if

(a) there is an operation called addition and denoted by the symbol "+" with respect to which R is an abelian group and

(b) multiplication of functions f, g in R by complex numbers, α, β, is defined so that

(1) $\alpha(f + g) = \alpha f + \alpha g$

(2) $(\alpha + \beta)f = \alpha f + \beta f$

(3) $\alpha(\beta f) = (\alpha\beta)f$

(4) $1 \cdot f = f$

(5) $0 \cdot f = 0$

Definition 2 The scalar product (inner product) on a linear space R is a mapping from R x R onto C, denoted by (f, g), such that

(a) $(g, f) = \overline{(f, g)}$, where the bar denotes conjugation.

(b) $(\alpha_1 f_1 + \alpha_2 f_2, g) = \alpha_1 (f_1, g) + \alpha_2 (f_2, g)$

(c) $(f, f) \geq 0$, with equality only if $f = 0$.

In this case, R is called an inner product space.

Definition 3 A linear space R is a metric space if there exists a function, ρ, defined on R x R such that

(a) $\rho(f, g) \geq 0$ and equality holds only if $f = g$.

(b) $\rho(f, g) = \rho(g, f)$

(c) $\rho(f, g) \leq \rho(f, h) + \rho(h, g)$ for any h in R.

Definition 4 A linear space R is said to be normed if for each f in R there is a real number, denoted by $||f||$, called the norm of f, such that

(a) $||f|| \geq 0$

(b) $||f|| = 0$ implies $f = 0$

(c) $||\alpha f|| = |\alpha| \ ||f||$

(d) $||f_1 + f_2|| \leq ||f_1|| + ||f_2||$ (triangle inequality)

Definition 5 A Hilbert space, H, is an inner product space in which the norm of an element x is defined by

$$||x|| = \sqrt{(x,x)},$$

and a distance function, ρ, is defined by

$$\rho(x,y) = ||x - y||,$$

so that H becomes a metric space. In the following, the letter H will be used exclusively to denote a Hilbert space.

Definition 6 Two functions f and g in an inner product space are orthogonal if

$$(f, g) = 0$$

Theorem 2.1 Schwarz's inequality holds for all functions in H, namely

$$|(f, g)| \leq ||f|| \cdot ||g||.$$

Definition 7 For the functions , $\{Z_1\ Z_2, \ldots, Z_n\}$, the following determinant is known as the Gramian, or Gram determinant, of the functions:

$$\begin{vmatrix} (Z_1, Z_1) & (Z_1, Z_2) & \cdots\cdots\cdots & (Z_1, Z_n) \\ (Z_2, Z_1) & (Z_2, Z_2) & \cdots\cdots\cdots & (Z_2, Z_n) \\ \cdots & \cdots & \cdots & \cdots \\ (Z_n, Z_1) & (Z_n, Z_2) & \cdots\cdots\cdots & (Z_n, Z_n) \end{vmatrix}$$

Theorem 2.2 A necessary and sufficient condition for the linear independence of a set of functions is the non-vanishing of their Gram determinant.

Definition 9 The sequence of elements $\{x_n\}$ in H is said to converge strongly to x if $x \in H$ and

$$||x_n - x|| \to 0, \; n \to \infty$$

The element, x, is called the limit of the sequence. The convergence of the sequence $\{x_n\}$ to x will be denoted by

$$x_n \to x.$$

Theorem 2.3 In a Hilbert space, if $x_n \to x$ then

$$(x_n, y) \to (x, y)$$

for any function $y \in H$, and

$$||x_n|| \to ||x||.$$

Definition 10 A sequence $\{x_k\}$ in a Hilbert space is called a fundamental sequence if for any given $\epsilon > 0$ there exists a number N such that for $m, n > N$

$$||x_n - x_m|| < \epsilon.$$

Theorem 2.4 If $x_n \to x$ then $\{x_n\}$ is a fundamental sequence.

Definition 11 A Hilbert space is complete if every fundamental sequence converges to an element of the space. Unless otherwise stated, we will consider only complete Hilbert spaces.

Theorem 2.5 Any metric space can be embedded in a complete metric space; in particular, any Hilbert space can be embedded in a complete Hilbert space. The new space is called the closure or completion of the old space.

Definition 12 If the Hilbert space H' is not complete and is embedded in the complete Hilbert space H, and if the sequence $\{x_k\}$ of elements of H' is fundamental but does not converge to an element of H', its limit in H will be called an ideal element of the sequence $\{x_k\}$.

Definition 13 A subset L of a linear space R is called a linear manifold if f, g $\in$ L implies that $\alpha f + \beta g \in L$ for arbitrary numbers α, β.

Definition 14 If a linear manifold contains all its ideal elements; i.e, is closed, then it is called a subspace.

Theorem 2.5 If L is a subspace of the Hilbert space H then every element x in H can be represented uniquely in the form

$$x = y + z$$

where $y \in L$ and z is orthogonal to every element of L, $z \perp L$.

The element y is called the orthogonal projection of x on L.

Definition 15 An operator, A, on H is a correspondence that assigns to every element x in H another specific element y in H, denoted

$$y = Ax$$

Definition 16 The operator A is linear if $\alpha f + \beta g \in H$ implies

$$A(\alpha f + \beta g) = \alpha Af + \beta\, Ag.$$

Definition 17 The operator A is bounded if there exists a positive number M such that

$$||A\,x|| \leq M\,||x||$$

for all $x \in H$. The smallest such M is called the norm of A and is denoted by $||A||$.

Definition 18 The operations of addition and multiplication of bounded linear operators are defined as follows

$$(\alpha A)\,x = \alpha(Ax)$$

$$(A + B)\,x = Ax + Bx$$

$$(AB)\,x = A(Bx)$$

Definition 19 If A is a bounded linear operator, the adjoint operator, A*, is defined by the requirement that

$$(Ax,\ y) = (x,\ A^*y)$$

for all x, y in the space. The operator A is called self-adjoint if $A = A^*$.

Theorem 2.7 For bounded linear operators the following are true

(a) $||\alpha A|| = |\alpha|\ ||A||$

(b) $||A + B|| \leq ||A|| + ||B||$

(c) $||AB|| \leq ||A||\ ||B||$

(d) $(A^*)^* = A$

(e) $||A|| = ||A^*||$

Definition 20 An operator A is continuous if $x_n \to x$ implies $Ax_n \to Ax$.

Theorem 2.8 Every bounded linear operator is continuous, and every continuous linear operator is bounded.

Definition 21 A linear operator A has a bounded inverse if there exists a bounded operator B such that

$$AB = BA = I$$

where I is the identity operator.

Definition 22 The following is called an equation of the first kind

$$Ax = y,$$

and if A has a bounded inverse, A^{-1}, then the solution is given by

$$x = A^{-1}y.$$

Definition 23 The problem of finding a λ such that

$$Ax = \lambda x$$

for some $x \neq 0$ is called an eigenvalue problem. The number λ is called an eigenvalue of A and the corresponding functions, x, are called eigenfunctions of A corresponding to λ.

Definition 25 A sequence $\{A_k\}$ of bounded linear operators is said to converge strongly to the operator A, denoted by

$$A_n \to A$$

if for every $x \in H$, the sequence $\{A_n x\}$ converges to Ax.

Definition 26 A sequence $\{A_n\}$ of bounded linear operators is said to converge uniformly to A if

$$\lim_{n \to \infty} ||A - A_n|| = 0.$$

This is also known as norm convergence.

Definition 27 If L is subspace of H, then the operator mapping any $x \in H$ onto its orthogonal projection in L is called the orthogonal projection onto the space L and will be denoted by E_L.

Theorem 2.9 The operator E_L has the following properties

(a) $||E_L|| = 1$

(b) $E_L^2 = E_L$

(c) $E_L^* = E_L$

(d) $||I - E_n|| = 1$

Definition 28 If A is an operator on H and L is a subspace of H with the property that $x \in L$ implies $Ax \in L$, then L is said to *reduce* the operator A.

Definition 29 Suppose the operator A' is defined on the subspace L of H and A is an operator on H such that $Ax = A'x$ for x in L. Then A is called an *extension* of A'. The operator A' is called the *restriction* of A to L.

Definition 30 An operator A is *completely continuous* if and only if every infinite sequence of functions, $\{x_k\}$, whose norms are uniformly bounded contains a sequence, $\{x_{k_n}\}$, for which the sequence $\{Ax_{k_n}\}$ is convergent.

Theorem 2.10 A completely continuous operator is bounded.

Definition 31 The following is an *equation of the second kind*

$$x = uAx + f,$$

where u is a number, f is a prescribed function in H, and A is an operator on H.

Theorem 2.11 For an equation of the second kindin which the operator A is completely continuous, the *Fredholm alternative* is valid, namely: either the equation has a unique solution for each prescribed f, or it cannot be solved for a given f and the associated homogeneous equation ($f = 0$) then has non-trivial solutions. The first part of the alternative holds if u is a regular value, and the second if u is a reciprocal eigenvalue. If u is a regular value, there exists a bounded inverse operator, $(I - uA)^{-1}$, and the unique solution is given by

$$x = (I - uA)^{-1}f.$$

Definition 32 The bounds of a self-adjoint operator, A, are, respectively, the smallest number M and the largest number m for which

$$m(x,x) \leq (Ax,x) \leq M(x,x)$$

for every $x \in H$. If $m > 0$, then A is positive definite.

Definition 33 The elements of the sequence, $\{P_1, P_2, \ldots\}$, in H are mutually orthogonal if

$$(P_j, P_k) = 0, \; j \neq k$$

and if, in addition,

$$(P_j, P_k) = 1, \; j = k$$

then the sequence is called orthonormal.

Theorem 2.12 The eigenvalues of a self-adjoint operator A are real numbers.

Definition 34 If A is an operator on H and Z_0 in an element of H, then the linear monifold $L_{A_{Z_0}}$ is defined as

$$L_{A_{Z_0}} = \{y \in H : y = Q(A)Z_0\},$$

where $Q(t)$ is any arbitrary poloynomial with real coefficients.

The linear manifold $L_{A_{Z_0}}$ will also be denoted by L_Z. The subspace $\overline{L_Z}$ will be denoted by H_Z.

Definition 35 If A is an operator on H, Z_0 is an element of H and m is a positive integer, then the linear manifold $L_{A_{Z_0}}(m)$ is defined as

$$L_{A_{Z_0}}(m) = \{y \in H : y = Qm(A)Z_0\}$$

where Qm(t) is an arbitrary polynomial of degree m with real coefficients. The closure of $L_{A_{Z_0}}(m)$ is a subspace of H and will be denoted by $H_{A_{Z_0}}(m)$.

Definition 36 The statement that the sequence $\{x_k\}$ is q - convergen to x means that there is a number N such that for $n > N$,

$$||x - x_n|| < q^{n+1}.$$

for any previously assigned number $q>0$. Similarly, we say a sequence $\{\alpha_n\}$ of numbers is q - convergent to a if

$$|a - \alpha_n| < q^{n+1}.$$

Theorem 2.13 Weierstrass Approximation Theorem): If f is a continuous function on the non-degenerate interval I, then there exists a sequence $\{P_n(t)\}$ of polynomials converging uniformly to f on I.

Theorem 2.14 (Rouche's Theorem) Suppose f is analytic in the region R, which is bounded by the simple closed curve C and f can be written as

$$f(Z) = g(Z) + h(Z),$$

where $|g(Z)| > |h(Z)|$ on C.

Then the number of zeros of g in R is the same as the number of zeros of f in R.

Definition 37 If the operator A is defined on a subset K of H and for all x and y in K

$$(Ax, y) = (x, Ay)$$

then A is called a symmetric operator.

Theorem 2.15 If A is a symmetric operator defined on all of H then A is self-adjoint.

Theorem 2.16 If A is a positive definite operator then A has a bounded inverse.

Definition 38 Suppose A_0 is a positive definite self-adjoint operator defined on a linear manifold L_A which is dense in H, then a new scalar product defined by

$$[x,y] = (A_0x, y)$$

converts L_A into a Hilbert space. The closure of this space will be denoted by H_0.

Definition 39 If A is an operator on a subset K of H we define $\bar{A}$, the closure of A, as follows: for $x \in K$, $\bar{A}x = Ax$, and for x an ideal element of K corresponding to the sequence $\{x_n\}$, we define

$$\bar{A}x = \lim_{n \to \infty} Ax_n$$

Definition 40 If A is an operator then the set of all complex numbers α such that $(A - \alpha I)^{-1}$ fails to exist as a bounded operator defined on H is the spectrum of A, and will be denoted by $\sigma(A)$.

CHAPTER III

THE METHOD OF MOMENTS IN HILBERT SPACE

It is the purpose of this section to define the moment problem, to show how its solution can be used to solve problems of the first and second kind and to approximate any arbitrary bounded linear operator.

A. The Problem of Moments

Suppose Z_0, $Z_1, \ldots Z_n$ are distinct linearly independent elements of a Hilbert space H. Denote by H_n the subspace generated by $\{Z_0, Z, \ldots, Z_{n-1}\}$. Then H_n has dimension n. The following problem will be called the problem of moments:

Find a linear operator A_n defined on H_n such that

$$\begin{aligned} A_n Z_0 &= Z_1 \\ A_n Z_1 &= Z_2 \\ &\cdots\cdots \\ A_n Z_{n-2} &= Z_{n-1} \\ A_n Z_{n-1} &= E_n Z_n \end{aligned} \tag{3.1}$$

where $E_n Z_n$ is the orthogonal projection of Z_n on H_n.

Actually, A_n, the solution to the moments problem is already formulated. This is clear since all that is necessary to completely describe any linear operator is to determine its action on a basis. This is exactly how A_n was defined.

B. Determination of Eigenvalues of A_n.

Suppose that λ is an eigenvalue and x is an eigenfunction of the operator A_n, that is,

$$A_n x = \lambda x. \tag{3.2}$$

Let $x = \beta_0 Z_0 + \beta_1 Z_1 + \ldots + \beta_{n-1}$, where the β_k are scalars. Substituting this expression for x in (3.2) yields

$$A_n \left(\sum_{k=0}^{n-1} \beta_k Z_k \right) = \lambda \left(\sum_{k=0}^{n-1} \beta_k Z_k \right),$$

or

$$\sum_{k=0}^{n-2} \beta_k Z_{k+1} + \beta_{n-1}(E_n Z_n) = \lambda \left(\sum_{k=0}^{n-1} \beta_k Z_k \right).$$

Let

$$E_n Z_n = - \sum_{j=0}^{n-1} \alpha_j Z_j. \tag{3.3}$$

Then,

$$\sum_{k=0}^{n-2} \beta_k Z_{k+1} - \beta_{n-1} \left(\sum_{k=0}^{n-1} \alpha_k Z_k \right) = \lambda \left(\sum_{k=0}^{n-1} \beta_k Z_k \right),$$

or, combining terms,

$$\alpha_0 \beta_{n-1} Z_0 + (\beta_0 - \alpha_1 \beta_{n-1}) Z_1 + \ldots + (\beta_{n-2} - \alpha_{n-1} \beta_{n-1}) Z_{n-1} = \lambda \left(\sum_{k=0}^{n-1} \beta_k Z_k \right.$$

Since the Z_k are linearly independent, coefficients of like terms can be equated to obtain the following system of equations for λ and the β_k:

$$\begin{aligned} -\alpha_0 \beta_{n-1} &= \lambda \beta_0 \\ \beta_0 - \alpha_1 \beta_{n-1} &= \lambda \beta_1 \\ \vdots \quad \cdots&\cdots\cdots \\ \beta_{n-2} - \alpha_{n-1} \beta_{n-1} &= \lambda \beta_{n-1}. \end{aligned} \tag{3.4}$$

In order for this system to have a non-trivial solution for the β_k, the determinant of the coefficients, $P_n(\lambda)$, must be equal to zero. Therefore, the eigenvalues of A_n are the roots of

$$P_n(\lambda) = \lambda^n + \alpha_{n-1} \lambda^{n-1} + \ldots + \alpha_0 = 0. \tag{3.5}$$

A system of equations can be obtained to determine the coefficients, α_k, in $P_n(\lambda)$ as follows:

From (3.1) and (3.3) it follows that

$$P_n(A_n)Z_0 = (A_n^n + \alpha_{n-1} A_n^{n-1} + \ldots + \alpha_0 I)\, Z_0 = 0 \quad (3.6)$$

The equation $(E_n Z_n, Z_j) = (Z_n, Z_j)$ becomes

$$\sum_{k=0}^{n-1} (Z_j, Z_k)\alpha_k + (Z_j, Z_n) = 0 \quad (3.7)$$

The determinant of the coefficients of (3.7) is the Gramian of $Z_0, \ldots Z_{n-1}$, and so does not vanish by virtue of their linear independence. Thus (3.7) has a unique solution. Throughout the following, the α_k will denote the solution of system (3.7).

The coefficient, β_{n-1}, in (3.4) must differ from zero, otherwise all β_k would vanish and yield only the trivial solution. Since an eigenfunction is determined only to within a constant multiple, we can set $\beta_{n-1} = 1$ and, using (3.4), find the other β_k recursively from:

$$\beta_j = \lambda\beta_{j+1} + \alpha_{j+1} \qquad j = 0, 1, \ldots, n-3$$

$$\beta_{n-2} = \lambda + \alpha_{n-1}. \quad (3.8)$$

C. Problems of the First Kind in H_n.

Theorem 3.1 Suppose

$$A_n x = f \quad (3.9)$$

where x and f are in H_n and f has the representation

$$f = \sum_{k=0}^{n-1} b_k Z_k.$$

Then, if zero is not an eigenvalue of A_n,

$$x = (b_1 - \frac{b_0}{\alpha_0}\alpha_1)Z_0 + \ldots + (b_{n-1} + \frac{b_0}{\alpha_0}\alpha_{n-1})Z_{n-2} - \frac{b_0}{\alpha_0} Z_{n-1}.$$
$$(3.10)$$

Proof:

Define

$$F_{n-1}(A_n)Z_0 = \sum_{k=0}^{n-1} b_k A_n^k Z_0 = f \qquad \text{and}$$

$$F_{n-1}(t) = b_{n-1} t^{n-1} + b_{n-2} t^{n-2} + \dots + b_0.$$

By taking the scalar product of f successively with $Z_0, Z_1, \dots, Z_{n-1}$ the following system is obtained for the coefficients of the polynomial $F_{n-1}(t)$:

$$\sum_{k=0}^{n-1} (Z_j, Z_k)b_k - (Z_j, f) = 0, \ j = 0, 1, \dots, n-1 \qquad (3.11)$$

Since x is in H_n, it may be expressed in terms of a polynomial of degree (n-1) in A_n,

$$x = Q_{n-1} (A_n)Z_0. \qquad (3.12)$$

The substitution of (3.12) into (3.9) yields

$$A_n[Q_{n-1}(A_n)Z_0] = F_{n-1}(A_n)Z_0 = f,$$

or

$$[A_n Q_{n-1}(A_n) - F_{n-1}(A_n)]Z_0 = 0.$$

Thus, $[A_n Q_{n-1}(A_n) - F_{n-1}(A_n)]$ is an n-th degree polynomial in A_n which annihilates Z_0. Therefore it can differ from any other n-th degree polynomial which annihilates Z_0 by only a constant factor, C. To see this, let

$$T_n(A_n)Z_0 = R_n(A_n)Z_0 = 0,$$

where

$$T_n(A_n) = t_n A_n^n + t_{n-1} A_n^{n-1} + \dots + t_0 I$$

and

$$R_n(A_n) = \gamma_n A_n^n + \gamma_{n-1} A_n^{n-1} + \dots + \gamma_0 I.$$

Thus,

$$\left[\left(\frac{t_{n-1}}{t_n} - \frac{\gamma_{n-1}}{\gamma_n}\right) A_n^{n-1} + \ldots + \left(\frac{t_0}{t_n} - \frac{\gamma_0}{\gamma_n}\right)\right] Z_0 = 0,$$

or

$$\left(\frac{t_{n-1}}{t_n} - \frac{\gamma_{n-1}}{\gamma_n}\right) Z_{n-1} + \ldots + \left(\frac{t_0}{t_n} - \frac{\gamma_0}{\gamma_n}\right) Z_0 = 0.$$

Hence, by the linear independence of the Z_k,

$$\frac{t_{n-1}}{t_n} = \frac{\gamma_{n-1}}{\gamma_n}, \ldots, \frac{t_0}{t_n} = \frac{\gamma_0}{\gamma_n},$$

or

$$\frac{\gamma_n}{t_n} t_{n-1} = \gamma_{n-1}, \ldots, \frac{\gamma_n}{t_n} t_0 = \gamma_0,$$

and it follows that T_n and R_n differ by a constant factor as claimed.

Therefore

$$A_n Q_{n-1}(A_n) - F_{n-1}(A_n) = CP_n(A_n), \tag{3.13}$$

and

$$C = \frac{-F_{n-1}(0)}{P_n(0)} = \frac{-b_0}{\alpha_0}. \tag{3.14}$$

Thus, combining (3.13) and (3.14)

$$Q_{n-1}(A_n) = A_n^{-1}\left[F_{n-1}(A_n) - \frac{b_0}{\alpha_0} P_n(A_n)\right],$$

and thus the required solution of (3.9) is

$$x = A_n^{-1}\left[F_{n-1}(A_n) - \frac{b_0}{\alpha_0} P_n(A_n)\right] z_0. \tag{3.15}$$

Substituting the above expression for $F_{n-1}(A_n)$ and $P_n(A_n)$ into (3.15),

$$x = A_n^{-1}\left[\sum_{k=0}^{n-1} b_k A_n^k - \frac{b_0}{\alpha_0}(A_n^n + \alpha_{n-1}A_n^{n-1} + \ldots + \alpha_0 I)\right] Z_0$$

$$x = A_n^{-1}\left[(b_0 + b_1 A_n + b_2 A_n^2 + \ldots + b_{n-1}A_n^{n-1}) - \frac{b_0}{\alpha_0}(A_n^n + \ldots + \alpha_0 I)\right] Z_0$$

$$x = A_n^{-1}\left[(b_1 - \frac{b_0}{\alpha_0}\alpha_1)A_n + \ldots + (b_{n-1} - \frac{b_0}{\alpha_0}\alpha_{n-1})A_n^{n-1} - \frac{b_0}{\alpha_0}A_n^n\right] Z_0$$

$$x = (b_1 - \frac{b_0}{\alpha_0}\alpha_1)Z_0 + \ldots + (b_{n-1} - \frac{b_0}{\alpha_0}\alpha_{n-1})Z_{n-2} - \frac{b_0}{\alpha_0}Z_{n-1}.$$

This completes the proof.

Notice that to solve problem (3.9) in H_n it is necessary to calculate the inner products, (Z_i, Z_j), to solve (3.11) by inverting an $n \times n$ matrix and to solve system (3.7) by using this same inverse.

D. Equations of the Second Kind in H_n.

Theorem 3.2 Suppose

$$x = \mu A_n x + f \tag{3.16}$$

where x and f are in H_n and μ is not a reciprocal eigenvalue of A_n. Then

$$x = a_{n-1}Z_{n-1} + a_{n-2}Z_{n-2} + \ldots + a_0 Z_0, \tag{3.17}$$

where

$$a_0 = b_0 - \frac{F_{n-1}(1/\mu)}{P_n(1/\mu)}\alpha_0,$$

$$a_k = \mu a_{k-1} + b_k - \frac{F_{n-1}(1/\mu)}{P_n(1/\mu)}\alpha_k, \quad k = 1,2,\ldots,n-1.$$

and where

$$F_{n-1}(A_n)Z_0 = f$$

and $P_n(t)$ is as in equation (3.5).

Proof:

Substituting $x = Q_{n-1}(A_n)Z_0$ and $f = F_{n-1}(A_n)Z_0$ into (3.16) yields

$$Q_{n-1}(A_n)Z_0 = \mu A_n Q_{n-1}(A_n)Z_0 + F_{n-1}(A_n)Z_0,$$

or

$$\left[Q_{n-1}(A_n) - \mu A_n Q_{n-1}(A_n) - F_{n-1}(A_n)\right] Z_0 = 0.$$

The bracketed expression can differ from $P_n(A_n)$ by only a constant multiple, C. Hence,

$$Q_{n-1}(A_n) - \mu A_n Q_{n-1}(A_n) - F_{n-1}(A_n) = CP_n(A_n),$$

and thus

$$(1 - \mu t)Q_{n-1}(t) - F_{n-1}(t) = CP_n(t). \qquad (3.18)$$

Letting $t = 1/\mu$ it is clear that

$$C = \frac{- F_{n-1}(1/\mu)}{P_n(1/\mu)} .$$

Therefore

$$Q_{n-1}(A_n) = (I - \mu A_n)^{-1} \left[F_{n-1}(A_n) - \frac{F_{n-1}(1/\mu)}{P_n(1/\mu)}\right] ,$$

and the solution to (3.16) is given by

$$x = (I - \mu A_n)^{-1} \left[F_{n-1}(A_n) - \frac{F_{n-1}(1/\mu)}{P_n(1/\mu)}\right] Z_0. \qquad (3.19)$$

The solution can be written in terms of the basis elements. The substitution of

$$Q_{n-1}(t) = \sum_{k=0}^{n-1} a_k t^k,$$

$$F_{n-1}(t) = \sum_{k=0}^{n-1} b_k t^k,$$

$$P_n(t) = \sum_{k=0}^{n} \alpha_k t^k$$

into (3.18) yields

$$(1 - \mu t) \sum_{k=0}^{n-1} a_k t^k - \sum_{k=0}^{n-1} b_k t^k = C \sum_{k=0}^{n} \alpha_k t^k,$$

or,

$$\sum_{k=0}^{n-1} a_k t^k - \mu \sum_{k=0}^{n-1} a_k t^{k+1} = \sum_{k=0}^{n-1} b_k t^k + C \sum_{k=0}^{n} \alpha_k t^k.$$

However

$$\sum_{k=0}^{n-1} a_k t^{k+1} = \sum_{k=1}^{n} a_{k-1} t^k,$$

and therefore

$$(a_0 - b_0 - C\alpha_0) + \sum_{k=1}^{n-1} (a_k - \mu a_{k-1} - b_k - C\alpha_k)t^k - (\mu a_{n-1} + C)t^n = 0.$$

If the coefficients of like powers of t are equated to zero and the value of C is inserted, the following recursion formulas for the coefficients a_k are obtained:

$$a_0 = b_0 - \frac{F_{n-1}(1/\mu)}{P_n(1/\mu)} \alpha_0,$$

$$a_k = \mu a_{k-1} + b_k - \frac{F_{n-1}(1/\mu)}{P_n(1/\mu)} \alpha_k, \quad k = 1,2,\ldots n-1.$$

The coefficient of t^n is zero by our choice of C. Thus,

$$x = Q_{n-1}(A_n)Z_0 = a_0 Z_0 + a_1 Z_1 + \ldots + a_{n-1} Z_{n-1},$$

with a_k as defined above. This proves the theorem.

Notice that to solve this problem in H_n it is necessary to calculate the inner products, (Z_i, Z_j), to solve systems (3.7) and (3.11) by inverting a single n×n matrix and, using the α_k and b_k thus determined, compute $F_{n-1}(1/\mu)$ and $P_n(1/\mu)$.

E. Approximation of Bounded Linear Operators.

In this section it will be shown that a sequence of operators may be selected by the method of moments so as to converge strongly to a preassigned bounded linear operator.

Suppose that A is a bounded linear operator in a Hilbert space H. Choosing an element $Z_0 \neq 0$, the following sequence of iterations is formed using the operator A:

$$Z_0,\ Z_1 = AZ_0;\ Z_2 = AZ_1 = A^2 Z_0, \ldots, Z_n = AZ_{n-1} = A^n Z_0, \ldots$$

which will be assumed to be linearly independent.

By solving the moment problem, a sequence of operators $\{A_n\}$

is determined, each defined on its own subspace H_n generated by $\{Z_0, Z_1, \ldots, Z_{n-1}\}$. The spaces increase in dimension with increasing n, and H_{n+1} contains H_n. Then (3.1) assumes the form

$$\begin{aligned} Z_k &= A^k Z_0 = A_n^k Z_0 \qquad k = 0,1,\ldots,n-1 \\ E_n Z_n &= E_n A^n Z_0 = A_n^n Z_0. \end{aligned} \tag{3.20}$$

<u>Theorem 3.3</u> For each n, $A_n x = E_n A E_n x$ for $x \in H_n$.

Proof:

Suppose $x \in H_n$. Then $x = \sum_{k=0}^{n-1} a_k Z_k$, and, by definition, $E_n x = x$.

Therefore

$$AE_n x = \sum_{k=0}^{n-1} a_k A Z_k,$$

or

$$AE_n x = \sum_{k=0}^{n-1} a_k A^{k+1} Z_0 = \sum_{k=0}^{n-2} a_k A_n^{k+1} Z_0 + a_{n-1} A^n Z_0.$$

Then

$$\begin{aligned} E_n A E_n x &= \sum_{k=0}^{n-2} a_k A_n^{k+1} Z_0 + a_{n-1} E_n A^n Z_0 \\ &= \sum_{k=0}^{n-1} a_k A_n^{k+1} Z_0 \\ &= \sum_{k=0}^{n-1} a_k A_n Z_k = A_n \left(\sum_{k=0}^{n-1} a_k Z_k \right) = A_n x, \end{aligned}$$

which was to be shown.

This theorem allows the domain of the operator A_n to be extended to the whole space H. That is, $E_n A E_n$ is defined on all of H, and its restriction to H_n is A_n.

<u>Corollary 3.1</u> The sequence of operators $\{A_n\}$ defined above is uniformly bounded with $||A_n|| \leq ||A||$.

Proof:

By Theorem 3.3, $A_n = E_n A E_n$, hence

$$||A_n|| = ||E_n A E_n|| \leq ||E_n|| \; ||A|| \; ||E_n|| = ||A||.$$

<u>Theorem 3.4</u> The subspace H_z (see Definition 34) reduces the operator A.

Proof:

Suppose x is in H_z. Then either x is in L_z or x is an ideal element of L_z. If x is in L_z then

$$x = P(A)Z_0$$

for some polynomial P(t) with real coefficients. Therefore

$$Ax = AP(A)Z_0$$

is in H_z because tP(t) is also a polynomial with real coefficients.

Suppose x is an ideal element of L_z. Then a sequence $\{x_n\}$ of elements of L_z exists so that

$$||x - x_n|| \to 0, \; n \to \infty.$$

But then

$$||Ax - Ax_n|| \le ||A|| \; ||x - x_n|| \to 0,$$

and since Ax_n is in L_z, this implies that Ax, the limit of elements of L_z, must belong to H_z. Thus, in either case, if x is in H_z then Ax is in H_z. Therefore, H_z reduces the operator A, and the theorem is proved.

<u>Theorem 3.5</u> If A is a bounded linear operator and $\{A_n\}$ a sequence of solutions of the problem of moments (3.20), then the sequence $\{A_n\}$ converges strongly to A in the subspace H_z.

Proof:

If $x = Q(A)Z_0 \in L_z$, then $Ax = A_n x$ for every n exceeding the degree of the polynomial Q(t) by two. To see this suppose

$$x = (a_k A^k + a_{k-1}A^{k-1} + \dots + a_0 I)Z_0$$

where

$n \geq k + 2$. Then

$$Ax = (a_k A^{k+1} + a_{k-1}A^k + \dots + a_0 A)Z_0$$

$$= a_k A_{k+1} + a_{k-1}Z_k + \dots + a_0 Z_1.$$

However,

$$Z_\gamma = A^\gamma Z_0 = A_n^\gamma Z_0, \ \gamma = 0, 1, \dots, n-1$$

and since $k \leq n-2$ it is assured that

$$Z_{k+1} = A_n^{k+1} Z_0, \ Z_k = A_n^k Z_0, \ \dots, Z_1 = A_n Z_0.$$

Hence $A_n x = a_k Z_{k+1} + a_{k-1}Z_k + \dots + a_0 Z_1$ also. That is $Ax = A_n x$.
So, for each $x \in L_Z$ there exists an integer m such that $A_m x = Ax$.

Thus as $n \to \infty$, $||A_n x - Ax|| \to 0$ for any $x \in L_Z$.

Now suppose x is an ideal element of L_Z and let $\{x_m\}$ be a sequence of elements of L_Z converging to x. Then

$$||Ax - A_n x|| \leq ||Ax - Ax_m|| + ||Ax_m - A_n x_m|| + ||A_n x_m - A_n x||$$

by the triangle inequality.

Thus, since $||A_n|| \leq ||A||$,

$$||Ax - A_n x|| \leq 2||A|| \ ||x - x_m|| + ||A_n x_m - Ax_m||.$$

The first term on the right approaches zero as m increases because $x_m \to x$, and since x_m is in L_Z, the second term is identically zero for sufficiently large n.

Hence, $||Ax - A_n x|| \to 0$, $n \to \infty$ and the theorem is proved.

Thus far it has been assumed that the elements, Z_k, are linearly independent. Theorem 3.5 remains valid in the linearly dependent case. To see this, suppose that Z_n is dependent on $\{Z_0, \ldots, Z_{n-1}\}$. By definition, $AZ_{n-1} = Z_n$ and $A_nZ_{n-1} = E_nZ_n$. But in this case $E_nZ_n = Z_n$, hence $AZ_{n-1} = A_nZ_{n-1}$. This implies that for any x in H_n, $Ax = A_nx$, or, in other words, H_n reduces the operator A. It is also clear that $Q(A)Z_0$ is in H_n for every polynomial Q(t) with real coefficients Also, $Z_k = G_k(A)Z_0$, $k = 0, 1, \ldots, n-1$, where $G_k(t)$ is a polynomial of degree k, namely, t^k. Since H_n is closed, $H_Z = H_n$. Theorem 3.5 states that $||Ax - A_kx|| \to 0, k \to \to \infty$, and in this case for $k \geq n$ we have $||Ax - A_kx|| = 0$.

Therefore, Theorem 3.5 remains valid.

CHAPTER IV

COMPLETELY CONTINUOUS OPERATORS

In this section, the method of moments will be used to solve problems of the second kind and the eigenvalue problem, each involving a completely continuous operator.

A. Approximation of a Completely Continuous Operator

Theorem 4.1 If A is a completely continuous operator, then the sequence of operators, $\{A_n\}$, solving the problem of moments, (3.20), converges in norm to A in the subspace H_Z;

$$||A - A_n|| \to 0, \; n \to \infty.$$

Proof:

The theorem will be proved by contradiction. Suppose $||A - A_n|| \not\to 0$, $n \to \infty$. Then, for each n, $||A - A_n|| > 0$. Let $q = \text{glb}\{||A - A_n||\}$

Then there exists a sequence, $\{f_n\}$, which may be chosen so that $||f_n|| = 1$ for each n, such that for arbitrarily large n,

$$||(A - A_n)f_n|| \geq q > 0.$$

Since A is completely continuous, a convergent subsequence, $\{Af_{n_j}\}$, may be chosen from the sequence $\{Af_n\}$. The sequence $\{(I - E_n)f_{n_j}\}$ is uniformly bounded, i.e., $||(I - E_n)f_{n_j}|| \leq ||f_{n_j}|| = 1$. Therefore, a convergent subsequence may again be selected from the sequence $\{||A(I - E)f_{n_j}||\}$. Let $\{f_n\}$ also denote this convergent subsequence and

$$g = \lim \{Af_n\},$$

$$h = \lim \{A(I - E_n)f_n\}.$$

Note that

$$A - A_n = A - E_nAE_n = (I - E_n)A + E_nA(I - E_n).$$

Since

$$||(I - E_n)Af_n|| \le ||(I - E_n)\ (Af_n-g)|| + ||(I - E_n)g||$$

$$\le ||Af_n-g|| + \qquad || g-E_ng|| \to 0$$

and

$$||E_nA(I - E_n)f_n|| \le ||A(I - E_n)f_n|| \to ||h||$$

and

$$||h||^2 = \lim[A(I - E_n)\ f_n,h] = \lim[f_n,(I - E_n)A^*h]$$

$$\le \lim||(I - E_n)A^*h|| = 0$$

it follows that

$$||(A - A_n)f_n|| \to 0$$

This is a contradiction and the theorem is proved.

B. Equations of the Second Kind

Consider the equation

$$x = \mu Ax + f \qquad (4.1)$$

with f a given element of H, A a completely continuous operator, and μ a number.

The Fredholm alternative [Theorem 2.11] applies in this case, and the following theorem of Banach holds.

<u>Theorem 4.2</u> (<u>Banach's Theorem</u>)

If $|\mu| < ||A||^{-1}$, then μ is a regular value for equation (4.1).

Proof:

Consider the series,

$$f + \mu Af + \mu^2A^2f + \ldots$$

and let $x^{(n)}$ denote its n-th partial sum.

Then since $|\mu|\ ||A|| < 1$,

$$||x^{(n)} - x^{(m)}|| = ||\mu^m A^m f + \mu^{m+1} A^{m+1} f + \dots + \mu^{n-1} A^{n-1} f||$$

$$\leq (|\mu|^m\ ||A||^m + |\mu|^{m+1}\ ||A||^{m+1} + \dots + |\mu|^{n-1}\ ||A||^{n-1})||f||$$

$$\to 0$$

when n and m approach infinity. But H is complete, and so the sequence, $\{x^{(n)}\}$, has a limit, x, in H.

Then,

$$\mu Ax = \mu Af + \mu^2 A^2 f + \dots = x - f$$

and hence

$$x = \mu Ax + f.$$

Note that Theorem 4.2 is valid for any bounded operator.

The procedure for solving (4.1) is to set $Z_0 = f$ and form the sequence of iterations

$$Z_1 = Af, \dots, Z_n = A_n f, \dots .$$

The solution of the approximate equation

$$x_n = \mu A_n x_n + f \tag{4.2}$$

is then

$$x_n = \sum_{k=0}^{n-1} a_k A^k Z_0$$

$$= a_0 f + a_1 Af + \dots + a_{n-1} A^{n-1} f,$$

where the a_k are determined recursively by means of the formulas

$$a_0 = 1 - \frac{\alpha_0}{P_n(1/\mu)},$$

$$a_k = \mu a_{k-1} - \frac{\alpha_k}{P_n\left(\frac{1}{\mu}\right)}, \quad k = 1, 2, \dots, n-1,$$

where $P_n(t)$ and α_k are as in (3.5) and (3.7).

The convergence of this method is assured by the next theorem.

Theorem 4.3 If a solution to (4.1) exists for each f in H and if A is completely continuous then for n sufficiently large equation (4.2) has a solution and the sequence $\{x_n\}$ converges to the solution of (4.1).

Proof:

Equation (4.2) can be expressed as

$$x_n - \mu A x_n + \mu(A - A_n)x_n = f.$$

By assumption, μ is a regular value of the equation, hence the Fredholm alternative insures the existence of the bounded inverse operator, $(I - \mu A)^{-1}$. Applying it to both sides of the equation yields

$$x_n - \mu(I - \mu A)^{-1}(A_n - A)x_n = (I - \mu A)^{-1}f = x \qquad (4.3)$$

According to Banach's theorem, a solution to (4.3) exists if

$$\|\mu(I - \mu A)^{-1}(A_n - A)\| < 1. \qquad (4.4)$$

But,

$$\|\mu(I - \mu A)^{-1}(A_n - A)\| \le |\mu| \, \|(I - \mu A)^{-1}\| \, \|A - A_n\|$$
$$\to 0, \; n \to \infty.$$

Therefore, condition (4.4) will be satisfied if n is sufficiently large. The existence of a solution to (4.2) is thus proved.

To prove the convergence, we start with the following representation of the inverse operator,

$$(I - \mu A_n)^{-1} = [I - \mu(I - \mu A)^{-1}(A_n - A)]^{-1}(I - \mu A)^{-1}.$$

Expanding in powers of μ, we obtain

$$[I - \mu(I - \mu A)^{-1}(A_n - A)]^{-1} = \sum_{k=0}^{\infty} \mu^k[(I - \mu A)^{-1}(A_n - A)]^k.$$

This is a convergent series since for sufficiently large n

$$||\mu(I - \mu A)^{-1} (A_n - A)|| < 1.$$

The sequence of operators, $\{(I - \mu A_n)^{-1}\}$, is uniformly bounded for a sufficiently large n since

$$||(I - \mu A_n)^{-1}|| \leq ||(I - \mu A)^{-1}|| \sum_{k=0}^{\infty} |\mu|^k \, ||(I - \mu A)^{-1}||^k \, ||A - A_n||^k$$

$$= \frac{||(I - \mu A)^{-1}||}{1 - |\mu| \, ||(I - \mu A)^{-1}|| . ||A - A_n||},$$

and by Theorem 4.1

$$||A - A_n|| \to 0, \; n \to \infty.$$

Therefore $||(I - \mu A_n)^{-1}|| \leq ||(I - \mu A)^{-1}||$

The sequence $\{(I - \mu A_n)^{-1}\}$ converges uniformly to $(I - \mu A)^{-1}$ in the subspace H_Z. To see this, consider the relation

$$(I - \mu A_n)^{-1} - (I - \mu A)^{-1} = \{[I - \mu (I - \mu A)^{-1}(A_n - A)]^{-1} - I\}(I - \mu A)^{-1}$$

$$= \sum_{k=1}^{\infty} \mu^k [(I - \mu A)^{-1} (A_n - A)]^k (I - \mu A)^{-1}.$$

Then

$$||(I - \mu A_n)^{-1} - (I - \mu A)^{-1}||$$

$$\leq ||(I - \mu A)^{-1}|| \sum_{k=1}^{\infty} |\mu|^k \, ||(I - \mu A)^{-1}||^k \, ||A_n - A||^k$$

$$= \frac{|\mu| \, ||(I - \mu A)^{-1}||^2}{1 - |\mu| \, ||(I - \mu A)^{-1}|| \, ||A - A_n||} \, ||A - A_n|| \to 0,$$

since $||A - A_n|| \to 0$ in H_Z as $n \to \infty$. Since f is in H_Z the sequence $\{(I - \mu A_n)^{-1} f\}$ converges to $x = (I - \mu A)^{-1} f$. This completes the proof of the theorem.

Notice that to solve (4.2) for fixed n, it is necessary to invert an n x n matrix from (3.7) to find the quantities α_k. As n increases, this becomes more difficult and, more important, less accurate since the determinant of the system decreases rapidly. It will now be shown how the method of moments can be used to speed the convergence of the well-known Liouville-Neumann method of successive approximations. One advantage will be that n is fixed throughout.

On the basis of the sequence of iterations

$$Z_0 = f,\ Z_, = Af, \ldots, Z_n = A^n f$$

and with fixed n we construct the operator A_n.

Equation (4.1) can be expressed as

$$x - \mu A_n x + \mu (A_n - A)x = f.$$

The operator $(I - \mu A_n)^{-1}$ exists when n is sufficiently large. Hence, applying it to both sides of the last equation, we obtain

$$x = \mu (I - \mu A_n)^{-1}(A - A_n)x + (I - \mu A_n)^{-1} f. \qquad (4.1')$$

Equation (4.1') will be solved by the Liouville-Neumann method. Define

$$x^0 = (I - \mu A_n)^{-1} f = \sum_{j=0}^{n-1} d_j^{(0)} Z_j$$

The element x^0 can be found by the methods of Chapter III. The successive approximations are obtained by the formula

$$x^{(k+1)} = \mu (I - \mu A_n)^{-1}(A - A_n)x^{(k)} + (I - \mu A_n)^{-1} f. \qquad (4.7)$$

To compute $x^{(k+1)}$ from $x^{(k)}$ it is required to solve the equation

$$x^{(k+1)} - \mu A_n x^{(k+1)} = \mu (A - A_n)x^{(k)} + f. \qquad (4.8)$$

Applying the operator E_n to both sides yields

$$E_n x^{(k+1)} - \mu A_n E_n x^{(k+1)} = \mu (E_n A - A_n)x^{(k)} + f. \qquad (4.9)$$

The solution $E_n x^{(k+1)}$ of this equation, denoted by $x_n^{(k+1)}$, can be found by the methods of Chapter III.

Substracting (4.9) from (4.8) yields

$$x^{(k+1)} - x_n^{(k+1)} = \mu(A-E_nA)x^{(k)},$$

or

$$x^{(k+1)} = x_n^{(k+1)} + \mu(A - E_nA)x^{(k)}. \tag{4.10}$$

In obtaining (4.9) and (4.10) we made use of the facts that $E_nf = f$ and

$$E_nA_n = E_n^2\,AE_n = A_n = E_nAE_n^2 = A_nE_n.$$

Assume that $x^{(k)}$ has the form

$$x^{(k)} = \sum_{j=0}^{n+k-1} d_j^{(k)} Z_j.$$

Then

$$(A - A_n)x^{(k)} = (A - E_nAE_n)x^{(k)}$$

$$= \sum_{j=0}^{n+k-1} d_j^{(k)}\, Z_{j+1} - \sum_{j=0}^{n-2} d_j^{(k)}\, Z_{j+1} - \sum_{j=n-1}^{n+k-1} d_j^{(k)}\, E_nAE_nZ_j$$

$$= \sum_{j=n-1}^{n+k-1} d_j^{(k)}\,(Z_{j+1} - E_nAE_nZ_j).$$

Now let

$$E_nZ_j = \sum_{s=0}^{n-1} r_s^{(j)}\, Z_s,$$

the $r_s^{(j)}$ being determined from the system of equations

$$(Z_p,\ Z_0)r_0^{(j)} + \ldots + (Z_p,\ Z_{n-1})r_{n-1}^{(j)} = (Z_p,\ Z_j)$$

$$p=0,\ 1, \ldots,\ n-1.$$

Then

$$AE_nZ_j = \sum_{s=0}^{n-1} r_s^{(j)}\, Z_{s+1},$$

and

$$E_nAE_nZ_j = \sum_{s=0}^{n-2} r_s^{(j)}Z_{s+1} + r_{n-1}^{(j)}\, E_nZ_n$$

$$= \sum_{s=1}^{n-1} r_{s-1}^{(j)} Z_s - r_{n-1}^{(j)} \sum_{s=0}^{n-1} \alpha_s Z_s$$

$$= \sum_{s=0}^{n-1} (r_{s-1}^{(j)} - r_{n-1}^{(j)} \alpha_s) Z_s. \tag{4.11}$$

$$r_{-1}^{(j)} = 0$$

Thus

$$(A - A_n)x^{(k)} = \sum_{j=n-1}^{n+k-1} d_j^{(k)} [Z_{j+1} - \sum_{s=0}^{n-1} (r_{s-1}^{(j)} - r_{n-1}^{(j)} \alpha_s) Z_s]$$

$$= \sum_{j=n-1}^{n+k-1} d_j^{(k)} (Z_{j+1} - E_n Z_{j+1})$$

$$+ \sum_{j=n-1}^{n+k-1} d_j^{(k)} [\sum_{s=0}^{n-1} (r_j^{(j+1)} - r_{s-1}^{(j)} + r_{n-1}^{(j)} \alpha_s) Z_s].$$

Applying E_n to both sides of this last equation yields

$$(E_n A - A_n)x^{(k)} = \sum_{j=n-1}^{n+k-1} d_j^{(k)} [\sum_{s=0}^{n-1} (r_s^{(j+1)} - r_{s-1}^{(j)} + r_{n-1}^{(j)} \alpha_s) Z_s]$$

$$= \sum_{s=0}^{n-1} Z_s [\sum_{j=n-1}^{n+k-1} d_j^{(k)} (r_s^{(j+1)} - r_{s-1}^{(j)} + r_{n-1}^{(j)} \alpha_s)]$$

and subtracting this from $(A - A_n)x^{(k)}$ above yields

$$(A - E_n A)x^{(k)} = \sum_{j=n-1}^{n+k-1} d_j^{(k)} (Z_{j+1} - E_n Z_{j+1})$$

$$= \sum_{j=n-1}^{n+k-1} d_j^{(k)} (Z_{j+1} - \sum_{s=0}^{n-1} r_s^{(j+1)} Z_s) = \sum_{s=0}^{n+k} h_s^{(k+1)} Z_s,$$

where

$$h_s^{(k+1)} = - \sum_{j=n-1}^{n+k-1} r_s^{(j+1)} d_j^{(k)} \qquad s=0, 1, \ldots, n-1$$

and

$$h_s^{(k+1)} = d_{s-1}^{(k)} \qquad s=n, \ldots, n+k$$

It still remains to solve (4.9),

$$x_n^{(k+1)} - \mu A_n x_n^{(k+1)} = \mu(E_n A - A_n)x^{(k)} + f. \qquad (4.12)$$

Putting,

$$x_n^{(k+1)} = x_0 + \mu \sum_{s=0}^{n-1} a_s^{(k+1)} Z_s,$$

we can find the coefficients $a_s^{(k+1)}$ by the methods of Chapter III. This yields

$$a_0^{(k+1)} = \sum_{j=n-1}^{n+k-1} d_j^{(k)} \left(r_0^{(j+1)} + r_{n-1}^{(j)} \alpha_0\right) - \frac{F_{n-1}^{(k+1)}\left(\frac{1}{\mu}\right)}{P_n\left(\frac{1}{\mu}\right)} \alpha_0$$

$$a_s^{(k+1)} = \sum_{j=n-1}^{n+k-1} d_j^{(k)} \left(r_s^{(j+1)} - r_{s-1}^{(j)} + r_{n-1}^{(j)} \alpha_s\right)$$

$$+ \mu a_{s-1}^{(k+1)} - \frac{F_{n-1}^{(k+1)}\left(\frac{1}{\mu}\right)}{P_n\left(\frac{1}{\mu}\right)} \alpha_s, \quad s=1, 2, \ldots, n-1$$

where

$$F_{n-1}^{(k+1)}(t) = \sum_{s=0}^{n-1} \left[\sum_{j=n-1}^{n+k-1} d_j^{(k)} \left(r_s^{(j+1)} - r_{s-1}^{(j)} + \alpha_s r_{n-1}^{(j)}\right) \right] t^s.$$

Thus, from (4.10) we have that

$$x^{(k+1)} = \sum_{s=0}^{n+k} d_s^{(k+1)} Z_s,$$

with

$$d_s^{(k+1)} = \mu h_s^{(k+1)} + d_s^{(0)} + \mu a_s^{(k+1)} \qquad s=0, 1, \ldots, n-1$$

$$d_s^{(k+1)} = \mu h_s^{(k+1)}. \qquad s=n, \ldots, n+k$$

In summary, we first must find $x^{(0)}$, which involves inverting the $n \times n$ matrix from (3.7). Finding the quantities, $r_s^{(j)}$, in (4.8) involves the same inverse. Finding the quantities, α_k, also involves this same inverse. Hence, this method has the advantage of keeping n fixed, so that computational errors introduced by increasing n are avoided.

The convergence of the method is assured because it is the Liou ville-Neumann method.

We now investigate the speed of convergence. Consider the basic formula

$$x^{(k+1)} = \mu(I - \mu A_n)^{-1}(A - A_n)x^{(k)} + (I - \mu A_n)^{-1} f. \quad (4.7)$$

Let B denote the operator

$$B = \mu(I - \mu A_n)^{-1} (A - A_n)$$

so that (4.7) can be written in the form

$$x^{(k+1)} = Bx^{(k)} + x^{(0)} \quad (4.7')$$

Since the operators $(I - \mu A_n)^{-1}$ are uniformly bounded for sufficien large n and since $||A - A_n|| \to 0$, given any arbitrarily small positive quantity, q, we can find an n such that

$$||B|| \le |\mu| \; ||(I - \mu A_n)^{-1}|| \; ||A - A_n|| < q.$$

Choose such a value of n,

From (4.7') it follows that

$$x^{(k)} = \sum_{j=0}^{k} B^j x^{(0)}.$$

If x denotes the exact solution of (4.1'), then according to Banach's Theorem, we have

$$x = \sum_{j=0}^{\infty} B^j x^{(0)}.$$

Subtraction then yields

$$||x - x^{(k)}|| = ||B^{k+1} \sum_{j=0}^{\infty} B^j x^{(0)}|| \le ||B||^{k+1} \; ||x||,$$

$$||x - x^{(k)}|| \le q^{k+1} \; ||x||. \tag{4.13}$$

Thus, the sequence $\{x^{(k)}\}$ is q-convergent to x.

Since $x^{(k)}$ is a polynomial in A of the form

$$x^{(k)} = G_{n+k-1}(A)f,$$

we have the following lemma.

<u>Lemma 4.1</u> If x is a solution of

$$x = \mu Ax + f$$

with μ a regular value and A a completely continuous operator, then there exists a sequence of polynomials $\{G_{m-1}(t)\}$ such that the sequence $\{G_{m-1}(A)Z_0\}$ is q-convergent to x.

<u>Proof</u>:

Consider $G_{m-1}(A)f = x^{(m-n)}$. Then inequality (4.13) implies that

$$||x - G_{m-1}(A)f|| = ||x - x^{(m-n)}|| \le q^{m-n+1} \; ||x|| = Cq^m,$$

where $C = ||x||q^{1-n}$ is independent of m. Thus $\{G_{m-1}(A)Z_0\}$ is q-convergent to x as claimed.

C. <u>The Eigenvalue Problem</u>

<u>Theorem 4.4</u> Suppose

$$x - \mu Ax = 0 \tag{4.14}$$

where A is a completely continuous operator and μ is a reciprocal eigenvalue; that is (4.14) has a non-trivial solution in H_Z. Then each such solution is unique, that is, to each eigenvalue there corresponds but one eigenfunction in H_Z. The sequences, $\{\mu_n\}$ and $\{x_n\}$, of solutions to

$$x_n - \mu_n A_n x_n = 0 \tag{4.15}$$

are q-convergent to the reciprocal eigenvalues and eigenfunctions, respectively.

Proof:

First, we express (4.14) in the form

$$x - \mu Ax = x - \mu(A - A_n)x - \mu A_n x = 0 \qquad (4.15')$$

By Banach's theorem, all values of μ in the circular region

$$|\mu| < \frac{1}{||A - A_n||} \qquad (4.16)$$

are regular for the operator $I - \mu(A - A_n)$ or, in other words, the inverse operator $[I - \mu(A - A_n)]^{-1}$ exists. The fulfillment of condition (4.16) can always be assured since with increasing n, $||A - A_n|| \to 0$.

Equation (4.14) may be expressed in the form

$$[I - \mu(A - A_n)]\{I - \mu[I - \mu(A - A_n)]^{-1}A_n\}x = 0,$$

which can be verified immediately by carrying out the indicated multiplication. Applying $[I - \mu(A - A_n)]^{-1}$ to both sides yields

$$x = \mu[I - \mu(A - A_n)]^{-1}A_n x \qquad (4.17)$$

Now, $A_n x \in H_n$ and hence is representable as

$$A_n x = \sum_{k=0}^{n-1} \frac{b_k}{\mu} Z_k$$

with certain constant coefficients b_k.

Let

$$\xi_k = [I - \mu(A - A_n)]^{-1} Z_k. \qquad (4.18)$$

Thus, from (4.17) and (4.18) it is seen that

$$x = \sum_{k=0}^{n-1} b_k \xi_k \qquad (4.19)$$

The ξ_k are linearly independent since if there exist numbers C_k not all zero such that

$$\sum_{k=0}^{n-1} C_k \xi_k = 0$$

then

$$\sum_{k=0}^{n-1} C_k \xi_k = [I - \mu(A-A_n)]^{-1} \sum_{k=0}^{n-1} C_k \xi_k = 0.$$

Hence, the last sum vanishes, contradicting the linear independence of the Z_k.

The equation defining ξ_k is equivalent to

$$\xi_k - \mu(A - A_n)\, \xi_k = Z_k.$$

Therefore,

$$\xi_k = Z_k,\ k = 0,\ 1,\ \ldots,\ n-2$$

and

$$A_n x = \sum_{k=0}^{n-2} b_k Z_{k+1} + b_{n-1} A_n \xi_{n-1}$$

$$[I - \mu(A - A_n)]^{-1} A_n x = \sum_{k=0}^{n-2} b_k \xi_{k+1} + b_{n-1}[I - \mu(A - A_n)]^{-1} A_n \xi_{n-1}. \tag{4.20}$$

Let

$$A_n \xi_{n-1} = - \sum_{s=0}^{n-1} \alpha_s(\mu) Z_s.$$

Then

$$[I - \mu(A - A_n)]^{-1} A_n \xi_{n-1} = - \sum_{s=0}^{n-1} \alpha_s(\mu) \xi_s. \tag{4.21}$$

From (4.17)

$$x = \mu[I - \mu(A - A_n)]^{-1} A_n x$$

substituting (4.20) into this yields

$$x = \mu\{ \sum_{k=0}^{n-2} b_k \xi_{k+1} + b_{n-1} [I - \mu(A - A_n)]^{-1} A_n \xi_{n-1}\}$$

then substituting (4.21) into this yields

$$x = \mu\left[\sum_{k=0}^{n-2} b_k\xi_{k+1} - b_{n-1}\sum_{s=0}^{n-1}\alpha_s(\mu)\xi_s\right].$$

Substituting (4.19) into this yields

$$\sum_{k=0}^{n-1} b_k\xi_k = \mu\left[\sum_{k=0}^{n-1} b_{k-1}\xi_k - b_{n-1}\sum_{k=0}^{n-1}\alpha_k(\mu)\xi_k\right]$$

and hence,

$$\sum_{k=0}^{n-1}\left[b_k - \mu b_{k-1} + \mu b_{n-1}\alpha_k(\mu)\right]\xi_{k=0}, \quad b_{-1} = 0$$

In view of the linear independence of the ξ_k,

$$b_k - \mu b_{k-1} + \mu\alpha_k(\mu)\, b_{n-1} = 0 \qquad k = 0, 1, \ldots, n-1 \quad (4.22)$$

Equating to zero the determinant of this system of homogeneous linear equations, we arrive at the following equation for the reciprocal eigenvalues, μ;

$$\Delta(\mu) = \frac{1}{\mu^n} + \frac{\alpha_{n-1}(\mu)}{\mu^{n-1}} + \ldots + \frac{\alpha_1(\mu)}{\mu} + \alpha_0(\mu) = 0 \quad (4.23)$$

After finding an eigenvalue, the corresponding eigenfunction is found by means of (4.6), the coefficients b_k being calculated by use of the recursion formula

$$b_k = \mu[b_{k-1} - \alpha_k(\mu)], \; k = 0, 1, \ldots, n-2 \quad b_{n-1} = 1 \quad (4.22')$$

But, since the b_k are uniquely determined, it follows that to each reciprocal eigenvalue there corresponds but one eigenfunction determined of course to within a constant multiple. This completes the first part of the proof.

For sufficiently large m $(m > n)$, the equation

$$x_m - \mu A_m x_m = 0$$

can be transformed as before into $x_m = \mu[I - \mu(A_n - A_n)]^{-1}A_n x.$ (4.17')

Similarly, we let $\xi_k^{(m)} = [I - \mu(A_m - A_n)]^{-1} Z_k$ and seek a olution x_m of the form:

$$x_m = \sum_{k=0}^{n-1} b_k^{(m)} \xi_k^{(m)} \ ; \ \xi_k^{(m)} = Z_k, \ k=0, 1, \ldots, n-2$$

This gives

$$b_k^{(m)} - \mu b_{k-1}^{(m)} + \mu \alpha_k^{(m)}(\mu) \, b_{n-1}^{(m)} = 0, \ k=0, 1, \ldots, n-1 \tag{4.22''}$$

where the quantities $\alpha_k^{(m)}(\mu)$ are determined by the relation

$$A_n \xi_{n-1}^{(m)} = - \sum_{s=0}^{n-1} \alpha_s^{(m)}(\mu) Z_s$$

The elements ξ_{n-1} and $\xi_{n-1}^{(m)}$ are respective solutions of

$$\xi_{n-1} - \mu(A - A_n)\xi_{n-1} = Z_{n-1}$$

$$\xi_{n-1}^{(m)} - \mu(A_m - A_n)\xi_{n-1}^{(m)} = Z_{n-1}.$$

Since the sequence $\{A_m\}$ converges in norm to A by Lemma 4.1, there exists a polynomial $\overline{Q}_s(A - A_n)Z_{n-1}$ such that

$$||\xi_{n-1} - \overline{Q_s}(A - A_n)Z_{n-1}|| \le C\, q^s$$

with q some positive number. But $\overline{Q}_s(A - A_n)Z_s = Q_{s+n-1}(A)Z_0$, so that for $m = n + s$

$$||\xi_{n-1} - Q_{m-1}(A)Z_0|| \le Cq^{m-n} = C_1 q^m \tag{4.24}$$

Now let $y_{n-1}^{(m)} = Q_{m-1}(A)Z_0$. Then

$$y_{n-1}^{(m)} - \mu(A_m - A_n)y_{n-1}^{(m)} = E_m\,[y_{n-1}^{(m)} - \mu(A - A_n)y_{n-1}^{(m)}],$$

and

$$\xi_{n-1}^{(m)} - \mu(A_m - A_n)\xi_{n-1}^{(m)} = Z_{n-1} = E_n[\xi_{n-1} - \mu(A - A_n)\xi_{n-1}].$$

Substracting these last two relations gives

$$\xi_{n-1}^{(m)} - y_{n-1}^{(m)} - \mu(A_m - A_n)\ (\xi_{n-1}^{(m)} - y_{n-1}^{(m)})$$
$$= E_n[\xi_{n-1} - y_{n-1}^{(m)} - \mu(A - A_n)(\xi_{n-1} - y_{n-1}^{(m)})].$$

Hence,

$$||\xi_{n-1}^{(m)} - y_{n-1}^{(m)}|| \leq ||[I - \mu(A_m - A_n)]^{-1}||\ ||I - \mu(A - A_n)||\ ||\xi_{n-1} - y_{n-1}^{(m)}||.$$

Thus, there exists a constant C_2 such that

$$||\xi_{n-1}^{(m)} - y_{n-1}^{(m)}|| \leq C_2 q^m$$

or, equivalently,

$$||y_{n-1}^{(m)} - \xi_{n-1}^{(m)}|| \leq C_2 q^m \tag{4.25}$$

Adding (4.24) and (4.25) and applying the triangle inequality yields

$$||\xi_{n-1} - \xi_{n-1}^{(m)}|| \leq C_3 q^m \tag{4.26}$$

Since

$$A_n \xi_{n-1} = -\sum_{k=0}^{n-1} \alpha_k(\mu) Z_k,$$

$$A_n \xi_{n-1}^{(m)} = -\sum_{k=0}^{n-1} \alpha_k^{(m)}(\mu) Z_k,$$

it follows that a similar kind of estimate holds for $\alpha_k(\mu)$ and $\alpha_k^{(m)}\ \mu$, namely,

$$|\alpha_k(\mu) - \alpha_k^{(m)}(\mu)| \leq C^{(k)} \cdot q^m$$

where $C^{(k)}$ is a constant independent of m.

Now let $\Delta_m(\mu)$ denote the characteristic polynomial for the operator A_m:

$$\Delta_m(\mu) = \frac{1}{\mu_n} + \frac{\alpha_{n-1}^{(m)}(\mu)}{\mu^{n-1}} + \ldots + \alpha_0^{(m)}(\mu)$$

From the estimate for $\alpha_k(\mu)$, it then follows that

$$|\Delta(\mu) - \Delta_m(\mu)| \leq C' q^m.$$

Suppose that μ_0 is a root of the equation $\Delta(\mu) = 0$ of multiplicity

ρ satisfying condition (3). Then for sufficiently small $\gamma = |\mu - \mu_0|$,

$$|\Delta(\mu)| \geq C'' \gamma^\rho.$$

For $\gamma = (C'/C'')^{1/\rho} q^{m/\rho}$, we further have

$$|\Delta(\mu) - \Delta_m(\mu)| \leq C''\gamma^\rho \leq |\Delta(\mu)|.$$

Therefore, by Rouche's Theorem, (Theorem 214), $\Delta_m(\mu) = \Delta(\mu) + \Delta_m(\mu) - \Delta(\mu)$ has exactly ρ roots $\mu_m^{(1)}, \ldots, \mu_m^{(\rho)}$ in the region $|\mu - \mu_0| \leq \gamma$ which satisfy the inequalities

$$|\mu_0 - \mu_m^{(k)}| \leq \bar{C}\,\bar{q}^m, \quad k = 1, \ldots \rho,$$

where

$$\bar{C} = \left(\frac{C'}{C''}\right)^{1/\rho}, \quad \bar{q} = q^{1/\rho}.$$

Hence, in view of the arbitrariness of q, the eigenvalues of the approximate equation (4.15) are q-convergent to the eigenvalues of (4.14). The estimate for the eigenvalues and the equations (4.22), (4.22') and (4.22") in turn imply the same sort of convergence for the eigenfunctions. This completes the proof of the theorem.

Thus, to obtain the approximate solutions, one must solve the equation (4.15). This can be done by the methods of Chapter III. Recall, however, that in order to determine the eigenvalues of A_n, it is necessary to invert an n x n matrix (3.7). This may be difficult for large values on n and may involve considerable loss of accuracy. The Liouville-Neumann method should be used in such cases. The quantities, $\alpha_k(\mu)$, involved are, in general, difficult to calculate. Keeping n fixed, it will now be shown how power series may be obtained for the coefficients $\alpha_k(\mu)$ in equations (4.22) and (4.23) determining the reciprocal eigenvalues and eigenfunctions.

The coefficients $\alpha_s(\mu)$ were defined by the relation

$$A_n\xi_{n-1} = -\sum_{s=0}^{n-1} \alpha_s(\mu)\, Z_s,$$

in which

$$\xi_{n-1} = [I - \mu(A - A_n)]^{-1}\, Z_{n-1} = \sum_{j=0}^{\infty} \mu^j (A - A_n)^j Z_{n-1}.$$

We first derive recursion formulas for the coefficients of the latter power series. Let

$$(A - A_n)^j Z_{n-1} = \sum_{s=0}^{n+j-1} d_s^{(j)} Z_s.$$

Then

$$(A - A_n)^{j+1} Z_{n-1} = \sum_{s=0}^{n+j-1} d_s^{(j)} Z_{s+1} - \sum_{s=0}^{n+j-1} d_s^{(j)}\, A_n Z_s$$

$$= \sum_{s=n-1}^{n+j-1} d_s^{(j)}\, Z_{s+1} - \sum_{s=n-1}^{n+j-1} d_s^{(j)}\, A_n Z_s.$$

From (4.11) we have

$$A_n Z_s = \sum_{k=0}^{n-1} (r_{k-1}^{(s)} - r_{n-1}^{(s)}\, \alpha_k) Z_k,$$

so that

$$(A - A_n)^{j+1} Z_{n-1} = \sum_{s=0}^{n+j} d_s^{(j+1)}\, Z_s$$

$$= \sum_{s=n-1}^{n+j-1} d_s^{(j)}\, Z_{s+1} - \sum_{k=0}^{n-1} Z_k \left[\sum_{s=n-1}^{n+j-1} d_s^{(j)}\, (r_{k-1}^{(s)} - r_{n-1}^{(s)}\, \alpha_k) \right].$$

Thus,

$$d_s^{(j+1)} = d_{s-1}^{(j)} \qquad s = n, \ldots, n+j$$

$$d_s^{(j+1)} = -\sum_{k=n-1}^{n+j-1} d_k^{(j)}\, (r_{s-1}^{(k)} - r_{n-1}^{(k)}\, \alpha_k) \qquad s=0, 1, \ldots, n-1$$

(4.24)

and

$$\xi_{n-1} = \sum_{j=0}^{\infty} \mu^j (A - A_n)^j Z_{n-1} = \sum_{j=0}^{\infty} \mu^j \sum_{s=0}^{n+j-1} d_s^{(j)} Z_s .$$

Next, $A_n \xi_{n-1}$ will be computed.

$$A_n \xi_{n-1} = \sum_{j=0}^{\infty} \mu^j [\sum_{s=0}^{n+j-1} d_s^{(j)} A_n Z_s]$$

$$= \sum_{j=0}^{\infty} [\sum_{s=0}^{n+j-1} d_s^{(j)} \sum_{k=0}^{n-1} (r_{k-1}^{(s)} - r_{n-1}^{(s)} \alpha_k) Z_k] \mu^j$$

$$= \sum_{k=0}^{n-1} Z_k \{ \sum_{j=0}^{\infty} \mu^j [\sum_{s=0}^{n+j-1} d_s^{(j)} (r_{k-1}^{(s)} - r_{n-1}^{(s)} \alpha_k)]\}$$

$$= - \sum_{k=0}^{n-1} \alpha_k(\mu) Z_k .$$

Therefore, the linear independence of the Z_k implies

$$\alpha_k(\mu) = - \sum_{j=0}^{\infty} [\sum_{s=0}^{n+j-1} d_s^{(j)} (r_{k-1}^{(s)} - r_{n-1}^{(s)} \alpha_k)]\mu^j \qquad (4.25)$$

Equation (4.25) expresses the coefficients $\alpha_k(\mu)$ as a power series. Recall equation (4.23),

$$\Delta(\mu) = 1/\mu^n + \frac{\alpha_{n-1}(\mu)}{\mu^{n-1}} + \ldots + \frac{\alpha_1(\mu)}{\mu} + \alpha_0(\mu) = 0,$$

for finding the reciprocal eigenvalues, μ. At first glance, it seems that (4.23) implies the existence of at most n reciprocal eigenvalues. This can not be the case, since we are working in an infinite dimensional space H. This problem is resolved by observing that the coefficients $\alpha_k(\mu)$ in (4.23) depend on μ also. Thus, it is true that each reciprocal eigenvalue, μ, satisfies some n-th degree polynomial but not all of them satisfy the same n-th degree polynomial.

Thus, to approximate the reciprocal eigenvalues of the operator A, first obtain an approximation for the quantities, $\alpha_k(\mu)$, by truncating the power series expression for the $\alpha_k(\mu)$, given in (4.25), at some integer γ. That is, let

$$\alpha_{k_\gamma}(\mu) = - \sum_{j=0}^{\gamma} \left[\sum_{s=0}^{n+j-1} d_s^{(j)} \left(r_{k-1}^{(s)} - r_{n-1}^{(s)} \alpha_k\right)\right] \mu^j. \quad (4.26$$

Then use (4.26) in (4.23) to determine the approximate reciprocal eigenvalues. Then, to obtain further, possibly more accurate approximations, one could increase γ and repeat the procedure.

After an approximate eigenvalue is found, say μ, then the corresponding approximate eigenfunction is found by use of (4.19) where the b_k are defined as follows

$$b_k = \mu[b_{k-1} - \alpha_{k_\gamma}(\mu)], \quad k=0, 1, \ldots, n-2$$

$$b_{n-1}=1, \; b_{-1} = 0$$

CHAPTER V

SELF-ADJOINT OPERATORS

A. The Eigenvalue Problem in a Finite Dimensional Space

Let A be a self-adjoint operator on a Hilbert space, H, of dimension n. Let $\{y_1, y_2, \ldots, y_n\}$ be an orthonormal basis for H. Consider the eigenvalue problem

$$Ax = \lambda x. \tag{5.1}$$

Let

$$x = a_1 y_1 + a_2 y_2 + \ldots + a_n y_n. \tag{5.2}$$

Taking the scalar product of (5.1) with the respective basis vectors yields the following system of linear homogeneous equations:

$$\begin{aligned}
&a_1(Ay_1, y_1) + a_2(Ay_2, y_1) + \ldots + a_n(Ay_n, y_1) = \lambda a_1 \\
&a_1(Ay_1, y_2) + a_2(Ay_2, y_2) + \ldots + a_n(Ay_n, y_2) = \lambda a_2 \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&a_1(Ay_1, y_n) + a_2(Ay_2, y_n) + \ldots + a_n(Ay_n, y_n) = \lambda a_n
\end{aligned} \tag{5.3}$$

Equating to zero the determinant of the system, we obtain the characteristic equation for the eigenvalues,

$$\begin{vmatrix}
(Ay_1, y_1) - \lambda & (Ay_2, y_1) & \ldots & (Ay_n, y_1) \\
(Ay_1, y_2) & (Ay_2, y_2) - \lambda & \ldots & (Ay_n, y_2) \\
\cdots & \cdots & \cdots & \cdots \\
(Ay_1, y_n) & (Ay_2, y_n) & \ldots & (Ay_n, y_n) - \lambda
\end{vmatrix} = 0 \tag{5.4}$$

After finding an eigenvalue, λ_k, the corresponding eigenfunction, x_k, can be found by solving (5.3) for the coefficients, a_k, to be used in (5.2). It follows that x_k is defined to within a constant multiple which may be chosen so that $||x_k|| = 1$.

Notice that the above process is valid for an arbitrary linear operator.

Eigenfunctions corresponding to distinct eigenvalues are mutually orthogonal. To see this, suppose

$$Ax_k = \lambda_k x_k$$
$$Ax_i = \lambda_i x_i.$$

Then

$$(Ax_k, x_i) = \lambda_k (x_k, x_i)$$
$$(Ax_i, x_k) = \lambda_i (x_i, x_k).$$

Hence, by the self-adjointness of A, we see that

$$0 = (Ax_k, x_i) - (Ax_i, x_k) = (\lambda_k x_k, x_i) - (\lambda_i x_i, x_k)$$
$$= (\lambda_k - \lambda_i)\ (x_k, x_i)$$

Therefore, when $\lambda_k \neq \lambda_i$, it follows that

$$(x_k, x_i) = 0.$$

Thus, the set of eigenfunctions of A can be taken to be an orthonormal basis for a subspace K of H, and then if $x \in K$ we have that

$$x = \sum_{i=1}^{k} (x, x_i)\ x_i.$$

Note that $k = n$ implies $K = H$. That is, if there are n distinct eigenvalues of A, then the set, $\{x_k\}_{k=1}^{n}$, is an orthonormal basis for H.

B. Spectral Theory for Self-Adjoint Operators

This section is important because it provides the tool by which convergence can be proved for several iterative methods which follow.

Suppose the operator A has n distinct eigenvalues, which we list in increasing order,

$$\lambda_1 < \lambda_2 < \dots < \lambda_n.$$

The spectral function of A, denoted by ξ_λ, is defined by

$$\xi_\lambda x = \begin{cases} 0 & \lambda < \lambda_1 \\ (x,\ x_1)\ x_1 & \lambda_1 \le \lambda < \lambda_2 \\ (x,\ x_1)\ x_1 + (x,\ x_2)\ x_2 & \lambda_2 \le \lambda < \lambda_3 \\ \dots\dots\dots\dots\dots\dots\dots\dots & \\ (x,\ x_1)\ x_1 + \dots + (x,\ x_{n-1})x_{n-1} & \lambda_{n-1} \le \lambda < \lambda_n \\ x & \lambda \ge \lambda_n \end{cases}$$

for any element x in H.

The real valued function $(\xi_\lambda x, x)$ is given by

$$(\xi_\lambda x,x) = \begin{cases} 0 & \lambda < \lambda_1 \\ |(x,\ x_1)| & \lambda_1 \le \lambda < \lambda_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots & \\ \sum\limits_{k=1}^{n-1} |(x,\ x_k)|^2 & \lambda_{n-1} \le \lambda < \lambda_n \\ ||x||^2 & \lambda \ge \lambda_n. \end{cases}$$

To see this, suppose we wish to calculate $(\xi_\lambda x,x)$ where $\lambda_k \le \lambda < \lambda_{k+1}$

Then

$$\begin{aligned}(\xi_\lambda x,x) &= [(x,x_1)x_1 + (x_1x_2)x_2 + \dots + (x,x_k)x_k,\ x] \\ &= [(x,x_1)x_1,x] + [(x_1x_2)x_2,x] + \dots + [(x,x_k)x_k,\ x] \\ &= (x,x_1)\ (x_1,x) + (x,x_2)\ (x_2,x) + \dots + (x,x_k)\ (x_k,\ x) \\ &= (x,x_1)\ \overline{(x,x_1)} + (x,x_2)\ \overline{(x,x_2)} + \dots + (x,x_k)\ \overline{(x,x_k)} \\ &= |(x,x_1)|^2 + |(x,x_2)|^2 + \dots + |(x,x_k)|^2 \\ &= \sum_{i=1}^{k} |(x,x_i)|^2\end{aligned}$$

Thus, $(\xi_\lambda x,x)$ is constant for those values of λ where A has no eigenvalues and increases by jumps at each eigenvalue, λ_k, by the amount, $|(x, x_k)|^2$. Now, let

$$\Delta_k \xi_\lambda x = (x,x_k)x_k.$$

Then the expression

$$Ax = \lambda_1(x, x_1)x_1 + \lambda_2(x,x_2)x_2 + \ldots + \lambda_n(x,x_n)x_n \tag{5.5}$$

becomes

$$Ax = \sum_{k=1}^{n} \lambda_k \Delta_k \xi_\lambda x.$$

Therefore, since $(\xi_\lambda x,x)$ is a step function with jumps at the λ_k, we have that

$$Ax = \int_{\lambda_1}^{\lambda_n} \lambda d\xi_\lambda x,$$

by definition of the Riemann - Stieltjes integral.

Applying A to both sides of (5.5) yields

$$A^2x = A[\lambda_1(x,x_1)x_1 + \lambda_2(x,x_2)x_2 + \ldots + \lambda_n(x,x_n)x_n]$$

$$= \lambda_1(x,x_1)Ax_1 + \lambda_2(x,x_2)Ax_2 + \ldots + \lambda_n(x,x_n)Ax_n$$

$$= \lambda_1^2 (x,x_1)x_1 + \lambda_2^2 (x,x_2)x_2 + \ldots + \lambda_n^2 (x,x_n)x_n.$$

Similarly

$$A^m x = \lambda_1^m (x,x_1)x_1 + \lambda_2^m (x,x_2)x_2 + \ldots + \lambda_n^m (x,x_n)x_n.$$

This formula may be extended to any function of A by defining

$$f(A)x = f(\lambda_1) (x,x_1)x_1 + \ldots + f(\lambda_n)(x,x_n)x_n \tag{5.6}$$

$$= \sum_{k=1}^{n} f(\lambda_k)\Delta_k \xi_\lambda x.$$

Then

$$f_1(A)\ f_2(A)\ x = \sum_{k=1}^{n} f_1(\lambda_k)\ f_2(\lambda_k)\ \Delta_k \xi_\lambda x. \tag{5.7}$$

Notice that in the definition of f(A), only the values of f(t) at the points, λ_k, are involved and so g(A) = f(A) whenever $g(\lambda_k) = f(\lambda_k)$, for each eigenvalue λ_k.

In a similar way, a continuous function of any bounded self-adjoint operator, B, can be defined by

$$f(B) = \int_m^M f(\lambda)\, d\xi_\lambda ; \tag{5.8}$$

that is, for $x \in H$,

$$f(B)x = \int_m^M f(\lambda)\, d\xi_\lambda x,$$

where m and M are the bounds of the operator,and in which the only values of f that effect the result are those at the points of the spectrum of the operator, where $d\xi_\lambda \neq 0$. Also, we have

$$f_1(A)\, f_2(A) = \int_m^M f_1(\lambda)\, f_2(\lambda)\, d\xi_\lambda . \tag{5.9}$$

The norm of the operator f(A) can be estimated as follows. Suppose $|f(\lambda)| \leq C$. Then

$$\begin{aligned} ||f(A)x|| &= || \int_m^M f(\lambda) d\xi_\lambda x|| \leq || \int_m^M |f(\lambda)|\, d\xi_\lambda x|| \\ &\leq || \int_m^M C d\xi_\lambda x|| \\ &= C\, || \int_m^M d\xi_\lambda x|| = C||x||, \end{aligned}$$

and therefore

$$||f(A)|| \leq C. \tag{5.10}$$

Notice that if A is a bounded self-adjoint operator which has an inverse, that the inverse operator can be expressed as

$$A^{-1} = \int_m^M \frac{1}{\lambda}\, d\xi_\lambda . \tag{5.11}$$

This is clear since

$$AA^{-1}x = \int_m^M \frac{\lambda}{\lambda}\, d\xi_\lambda x = x.$$

Theorem 5.1 The bounded self-adjoint operator A has a bounded inverse if and only if there exists an interval $[-\sigma, \sigma]$ containing no part of the spectrum of A.

Proof:

Suppose the spectrum of A lies outside the interval $[-\sigma, \sigma]$. Then since zero is not an eigenvalue of A, the inverse operator, A^{-1}, exists. If λ is an eigenvalue of A, it follows that

$$|\lambda^{-1}| \leq \sigma^{-1}.$$

To see this, suppose

$$|\lambda^{-1}| > \sigma^{-1}.$$

Then either

$$\lambda^{-1} > \sigma^{-1} \text{ or } \lambda^{-1} < -\sigma^{-1},$$

which implies that

$$0 < \lambda < \sigma \text{ or } 0 > \lambda > -\sigma$$

However, in either case, these inequalities imply that

$$\lambda \in [-\sigma, \sigma],$$

which is a contradiction to the hypothesis that no part of the spectrum of A lies in $[-\sigma, \sigma]$.

Thus, if no part of the spectrum of A lies in $[-\sigma, \sigma]$, then (5.11) and (5.10) together imply

$$||A^{-1}|| \leq \sigma^{-1}.$$

We will now show that if A^{-1} is a bounded operator, then an interval $[-\sigma, \sigma]$ exists containing no part of the spectrum of A. This will be proved by contradiction.

Suppose that A^{-1} is a bounded operator and every given interval

$[-\sigma, \sigma]$ contains part of the spectrum of A. Thus, zero is a limit point of the spectrum of A. By (5.11) we can express A^{-1} as

$$A^{-1} = \int_m^M \frac{1}{\lambda} d\xi_\lambda,$$

which implies that

$$||A^{-1}|| = || \int_m^M \frac{1}{\lambda} d\xi_\lambda ||.$$

However, since $\lambda \to 0$,

$$|| \int_m^M \frac{1}{\lambda} d\xi_\lambda || \to \infty,$$

which implies that A^{-1} is not bounded. This is a contradiction, and therefore an interval $[-\sigma, \sigma]$ exists containing no part of the spectrum of A. This completes the proof of the theorem.

Theorem 5.2 If A is a bounded self-adjoint operator with inverse and $\{A_n\}$ a sequence of solutions of the moment problem, then the sequence of spectral functions $\{\xi_\lambda^{(n)}\}$ of the operators, A_n, converges strongly to the spectral function of A in the subspace H_Z.

Proof:

The proof is based on the fact that the sequence $\{A_n\}$ converges strongly to A. The first consequence of the strong convergence of these operators is the strong convergence of polynomials of A_n. That is, if P(t) is any polynomial, then

$$P(A_n) \to P(A), \ n \to \infty.$$

To see this , suppose

$$P(t) = \alpha_k t^k + \alpha_{k-1} t^{k-1} + \ldots + \alpha_1 t + \alpha_0$$

and consider

$$\lim_{n \to \infty} P(A_n) = \lim_{n \to \infty} (\alpha_k A_n^k + \alpha_{k-1} A_n^{k-1} + \ldots + \alpha_1 A_n + \alpha_0)$$

$$= \lim_{n \to \infty} (\alpha_k A_n^k) + \lim_{n \to \infty} (\alpha_{k-1} A_n^{k-1}) + \ldots + \lim_{n \to \infty} (\alpha_1 A_n) + \lim_{n \to \infty} (\alpha_0)$$

$$= \alpha_k \lim_{n \to \infty} (A_n^k) + \alpha_{k-1} \lim_{n \to \infty} (A_n^{k-1}) + \ldots + \alpha_1 \lim_{n \to \infty} A_n + \alpha_0$$

$$= \alpha_k (\lim_{n \to \infty} A_n)^k + \alpha_{k-1} (\lim_{n \to \infty} A_n)^{k-1} + \ldots + \alpha_1 A + \alpha_0$$

$$= \alpha_k A^k + \alpha_{k-1} A^{k-1} + \ldots + \alpha_1 A + \alpha_0 = P(A).$$

By means of the Weierstrass Approximation Theorem this result may be extended to any function $p(\mu)$ continuous on the interval $|\mu| \leq ||A||$.

In particular, let

$$\rho(\mu) = e_\lambda(\mu - \lambda),$$

where

$$e_\lambda(\mu) = \begin{cases} 1 \text{ for } \mu \leq \lambda \\ 0 \text{ for } \mu > \lambda. \end{cases}$$

Then, $\rho(A_n) = \xi_\lambda^{(n)} A_n$,

and

$$\rho(A) = \xi_\lambda A,$$

and since $\rho(\mu)$ is continuous,

$$\xi_\lambda^{(n)} A_n \to \xi_\lambda A.$$

Now

$$||\xi_\lambda^{(n)} Ax = \xi_\lambda Ax|| \leq ||\xi_\lambda^{(n)} (A - A_n)x|| + ||\xi_\lambda^{(n)} A_n x - \xi_\lambda Ax||$$

$$\leq ||(A - A_n)x|| + ||\xi_\lambda^{(n)} A_n x - \xi_\lambda Ax|| \to 0,$$

and this implies convergence in H_Z.

Also

$$\xi_\lambda^{(n)} x = \xi_\lambda^{(n)} AA^{-1} x \to \xi_\lambda AA^{-1} x = \xi_\lambda x$$

for each x in the domain of A^{-1}. Since the domain of A^{-1} is dense in

H and the spectral functions, $\xi_\lambda^{(n)}$ are uniformly bounded, it follows that

$$\xi_\lambda^{(n)} \to \xi_\lambda$$

in H_Z as claimed.

C. The Eigenvalue Problems

Now consider the eigenvalue problem,

$$Ax = \lambda x, \tag{5.12}$$

where A is a bounded self-adjoint operator. To solve (5.12) by the method of moments select an arbitrary $Z_0 \in H$ and form the sequence of iterations,

$$Z_1 = AZ_0$$

$$Z_2 = A^2 Z_0$$

$$\cdots\cdots\cdots$$

$$Z_n = A^n Z_0$$

$$\cdots\cdots\cdots$$

Then for fixed n, solve the corresponding eigenvalue problem,

$$A_n x_n = \lambda_n x_n. \tag{5.12'}$$

The convergence of this method is assured by Theorem 5.2.

Notice that problem (5.12') can be solved by the methods of Chapter III. However, since $A_n = E_n A E_n$, we see that A_n is self-adjoint, and it will now be shown how the self-adjointness of A_n simplifies the solution of (5.12').

Let $\rho_0 = Z_0$

and iteratively

$$\rho_{k+1} = (A - a_k I)\rho_k - b_{k-1}\, \rho_{k-1}$$

where

$$a_k = (A\rho_k,\rho_k) \;/\; (\rho_k,\rho_k) = (A\rho_{k_1}\rho_k) \;/\; ||\rho_k||^2$$

and

$$b_{k-1} = ||\rho_k||^2 \;/\; ||\rho_{k-1}||^2$$

The elements, ρ_k, are mutually orthogonal, i.e., $(\rho_j,\rho_k)= 0, j\neq k$, and

$$\rho_k = \bar{P}_k(A)Z_0 \tag{5.13}$$

with $\overline{P}_k(A)$ a k-th degree polynomial in A. The polynomials $\overline{P}_k(t)$ each have leading coefficient unity and satisfy the recursion relations:

$$\overline{P}_{k+1}(t) = (t - a_k)\,\overline{P}_k(t) - b_{k-1}\,\overline{P}_{k-1}(t) \tag{5.14}$$

$$b_{-1} = 0,\; P_0 \equiv 1.$$

Notice that $\rho_k = \overline{P}_k(A)Z_0 = \overline{P}_k(A_n)Z_0 \in H_n$ for $k = 0, 1, \ldots, n-1$. Thus the elements, $\{\rho_0, \rho_1, \ldots, \rho_{n-1}\}$, form an orthogonal set in H_n. It is clear that

$$\overline{P}_n(A_n)Z_0 = E_n\overline{P}_n(A)Z_0 = E_n\rho_n,$$

but since ρ_n is orthogonal to each element of the set $\{\rho_0, \ldots, \rho_{n-1}\}$, which is a complete orthogonal set in the subspace H_n,

$$E_n\rho_n = (\rho_n, \rho_0)\frac{\rho_0}{||\rho_0||^2} + (\rho_n,\rho_1)\frac{\rho_1}{||\rho_1||^2}$$

$$+ \ldots + (\rho_n,\rho_{n-1})\frac{\rho_{n-1}}{||\rho_{n-1}||^2} = 0,$$

and therefore

$$\overline{P}_n(A_n)Z_0 = 0 \tag{5.15}$$

Thus, $\overline{P}_n(t)$ is an n-th degree polynomial annihilating Z_0 and can therefore differ from the polynomial $P_n(t)$ obtained in Chapter III by only a constant multiple. However, since $\overline{P}_n(t)$ and $P_n(t)$ have leading coefficient unity, they are equal. Thus, the roots of

$$\overline{P}_n(t) = 0$$

are the eigenvalues of the operator A_n.

The corresponding eigenfunctions could be obtained as in Chapter III, however they may be expressed directly in terms of $\rho_0, \rho_1, \ldots, \rho_{n-1}$.

Let

$$A_n x_k = \lambda_k x_k,$$

with λ_k a root of $\overline{P}_n(\lambda)$. We look for an x_k of the form

$$x_k = \sum_{j=0}^{n-1} \frac{(\mu_k, P_j)}{||\rho_j||^2} \rho_j.$$

Since A_n is self-adjoint, its eigenfunctions form a complete orthonormal set in H_n. Expanding ρ_j in terms of the eigenfunctions we obtain

$$\rho_j = \overline{P}_j(A_n)Z_0 = \sum_{k=0}^{n-1} \overline{P}_j(\lambda_k) (Z_0, x_k)x_k.$$

Therefore

$$(x_k, \rho_j) = (\mu_k, \sum_{k=0}^{n-1} \overline{P}_j (\lambda_k) (Z_0, x_k)x_k)$$

$$= \overline{P}_j(\lambda_k) (Z_0, \mu_k)$$

Thus, x_k can be expressed as

$$x_k = C \sum_{j=0}^{n-1} \overline{P}_j(\lambda_k)\rho_j$$

The constant C must be chosen so as to insure normalization. That is, we must choose C so that

$$||x_k||^2 = || C \sum_{j=0}^{n-1} \overline{P}_j(\lambda_k)\rho_j||^2 = 1$$

Therefore,

$$(C \sum_{j=0}^{n-1} \overline{P}_j(\lambda_k)\, \rho_j,\ C \sum_{j=0}^{n-1} \overline{P}_j(\lambda_k)\rho_j) = 1$$

$$= C^2 (\sum_{j=0}^{n-1} \overline{P}_j(\lambda_k)\rho_j,\ \sum_{j=0}^{n-1} \overline{P}_j(\lambda_k)\rho_j)$$

$$= C^2 \sum_{j=0}^{n-1} (\overline{P}_j(\lambda_k)\rho_j,\ \overline{P}_j(\lambda_k)\rho_j),$$

due to the orthogonality of the elements, ρ_k, and so

$$C^2 \sum_{j=0}^{n-1} \overline{P}_j^2\ (\lambda_k)\ ||\rho_j||^2 = 1$$

or

$$C= \frac{1}{\sqrt{\sum_{j=0}^{n-1} \overline{P}_j^2\ (\lambda_1)\ ||\rho_j||^2}}$$

and therefore

$$x_k = \frac{1}{\sqrt{\sum_{j=0}^{n-1} \overline{P}_j^2\ (\lambda_k)\ ||\rho_j||^2}} \cdot \sum_{j=0}^{n-1} \overline{P}_j(\lambda_k)\rho_j \tag{5.16}$$

Notice that in order to obtain the eigenvalues and corresponding eigenfunctions of A_n by the above method, the following steps must be taken: (1) Form the sequence, $\{Z_k\}$, (2) orthogonalize to find the sequence $\{\rho_k\}$, (3) determine all the polynomials, $\overline{P}_t(\lambda)$, $0 \le t \le n$, (4) find the eigenvalues as roots of $\overline{P}_n(\lambda) = 0$, (5) substitute each eigenvalue, λ_k, into equation (5.16) to find the corresponding eigenfunction. Observe that in contrast to the method presented in Chapter III, this method does not require inverting any matrix at all.

A procedure will now be developed by which the polynomials $\overline{P}_k(t)$ can be found up to and including $k = 2n-2$, without having

2

to increase n. Assume that the polynomials, $\overline{P}_k(t)$, have been determined up to and including $\overline{P}_n(t)$, and let

$$\overline{P}_n(t) = \sum_{s=0}^{n} \alpha_s^{(n)} t^s, \tag{5.17}$$

with $\alpha_n^{(n)} = 1$.

We first expand Z_0 in terms of the eigenfunctions of A_n,

$$Z_0 = \sum_{k=0}^{n-1} a_k x_k.$$

Consider the function

$$\varphi(t) = \frac{a_0^2}{t-\lambda_0} + \;\ldots\ldots + \frac{a_{n-1}^2}{t-\lambda_{n-1}} ,$$

where $\lambda_0, \lambda_1, \ldots, \lambda_{n-1}$ are the eigenvalues of A_n. We look for an expansion of $\varphi(t)$ in inverse powers of t of the form

$$\varphi(t) = \frac{c_0}{t} + \frac{c_1}{t^2} + \ldots + \frac{c_s}{t^{s+1}} + \ldots = \sum_{s=0}^{\infty} \frac{c_s}{t^{s+1}}. \tag{5.18}$$

We formally determine the coefficients in the series. We first write

$$\frac{a_k^2}{t-\lambda_k} = \frac{a_k^2}{t} \sum_{s=0}^{\infty} \left(\frac{\lambda_k}{t}\right)^s,$$

and hence

$$\varphi(t) = \sum_{k=0}^{n-1} \frac{a_k^2}{t-\lambda_k} = \sum_{s=0}^{\infty} \left[\sum_{k=0}^{n-1} a_k^2 \lambda_k^s \right] \frac{1}{t^{s+1}} .$$

Therefore

$$c_s = \sum_{k=0}^{n-1} a_k^2 \lambda_k^s .$$

Also we have that

$$Z_s = A_n^s Z_0 = \sum_{k=0}^{n-1} a_k \lambda_k^s x_k$$

and

$$(Z_s, Z_0) = \sum_{k=0}^{n-1} a_k^2 \lambda_k^s = c_s, \quad s=0,1,\ldots n-1.$$

In general, for any i and k,

$$c_{i+k} = (A_n^i Z_0,\ A_n^k Z_0),$$

since

$$(A_n^i Z_0, A_n^k Z_0) = (A_n^{i+k} Z_0, Z_0) = c_{i+k},$$

due to the self-adjointness of A_n.

The problem of moments also provides a way of computing the coefficients c_0 up to c_{2n-2} inclusive, namely

$$c_{i+k} = (Z_i, Z_k) \quad i+k \le 2n-2.$$

This is true because

$$(Z_i, Z_k) = (A_n^i Z_0, A_n^k Z_0) = c_{i+k}$$

for $0 \le i,k \le n-1$. Hence $0 \le i+k \le 2n-2$.

We now find $\overline{P}_n(t)$. Since the λ_k are its roots, and since its leading coefficient is unity, we can write

$$\varphi(t) = \sum_{k=0}^{n-1} \frac{a_k^2}{t-\lambda_k} = \frac{Q_{n-1}(t)}{\overline{P}_n(t)} = \frac{c_0}{t} + \frac{c_1}{t^2} + \ldots,$$

in which $Q_{n-1}(t)$ is some polynomial of degree (n-1).

Multiplying by $\overline{P}_n(t)$ yields

$$Q_{n-1}(t) = \bar{P}_n(t)\left(\frac{c_0}{t} + \frac{c_1}{t^2} + \ldots\right).$$

Substituting (5.17) yields

$$\sum_{s=0}^{n} \alpha_s^{(n)} t^s \cdot \sum_{j=0}^{\infty} \frac{c_j}{t^{j+1}} = \sum_{j=-n}^{\infty} \frac{1}{t^{j+1}} (c_j \alpha_0^{(n)} + \ldots + c_{j+n-1} \alpha_{n-1}^{(n)} + c_{j+n})$$
$$= Q_{n-1}(t).$$

Equating to zero the coefficients of negative powers of t we obtain

$$c_j \alpha_0^{(n)} + c_{j+1} \alpha_1^{(n)} + \ldots c_{j+n-1} \alpha_{n-1}^{(n)} + c_{j+n} = 0, \; j \ge 0.$$

Only n of these equations corresponding to j = 0,1,...n-1 are linearly independent, the others being linear combinations of them. This system corresponds exactly to system (3.7).

We now obtain recursion formulas for the coefficients of $\bar{P}_n(t)$. In the equality

$$Q_k(t) = \bar{P}_{k+1}(t)\left(\frac{c_0}{t} + \frac{c_1}{t^2} + \ldots + \frac{c_j}{t^{j+1}} + \ldots\right),$$

we replace $\bar{P}_{k+1}(t)$ by the recurrence relation obtained in (5.14) to obtain

$$Q_k(t) = (t - a_k)\,\bar{P}_k(t)\left(\frac{c_0}{t} + \frac{c_1}{t^2} + \ldots + \frac{c_k}{t^{k+1}} + \ldots\right)$$
$$- b_{k-1}\bar{P}_{k-1}(t)\left(\frac{c_0}{t} + \frac{c_1}{t^2} + \ldots + \frac{c_k}{t^{k+1}} + \ldots\right).$$

Now using $\bar{P}_k(t) = \sum_{j=0}^{k} \alpha_j^{(k)}\lambda^j$

and then equating to zero the coefficient of λ^{-k}, we obtain

$$c_k\alpha_0^{(k)} + c_{k+1}\alpha_1^{(k)} + \ldots + c_{2k-1}\alpha_{k-1}^{(k)} + c_{2k}$$
$$- b_{k-1}(c_{k-1}\alpha_0^{(k-1)} + c_k\alpha_1^{(k-1)} + \ldots + c_{2k-3}\alpha_{k-2}^{(k-1)} + c_{2k-2}) = 0$$

A similar operation of the coefficient of λ^{-k-1} yields

$$c_{k+1}\alpha_0^{(k)} + c_{k+2}\alpha_1^{(k)} + \ldots + c_{2k}\alpha_{k-1}^{(k)} + c_{2k+1}$$
$$- a_k(c_k\alpha_0^{(k)} + c_{k+1}\alpha_1^{(k)} + \ldots + c_{2k-1}\alpha_{k-1}^{(k)} + c_{2k})$$
$$- b_{k-1}(c_k\alpha_0^{(k-1)} + c_{k+1}\alpha_1^{(k-1)} + \ldots + c_{2k-2}\alpha_{k-2}^{(k-1)} + c_{2k-1}) =$$

Let

$$h_k = c_k\alpha_0^{(k)} + c_{k+1}\alpha_1^{(k)} + \ldots + c_{2k-1}\alpha_{k-1}^{(k)} + c_{2k},$$

$$h_{k+\frac{1}{2}} = c_{k+1}\alpha_0^{(k)} + c_{k+2}\alpha_1^{(k)} + \ldots + c_{2k}\alpha_{k-1}^{(k)} + c_{2k+1},$$

we have

$$b_{k-1} = \frac{h_k}{h_{k-1}}$$

$$a_k = \frac{h_{k+\frac{1}{2}}}{h_k} - b_{k-1}\frac{h_{k-\frac{1}{2}}}{h_k} = \frac{h_{k+\frac{1}{2}}}{h_k} - \frac{h_{k-\frac{1}{2}}}{h_{k-1}}$$

Thus the recurrence relation for $\bar{P}_{k+1}(t)$ shows

$$\bar{P}_{k+1}(t) = \sum_{s=0}^{k+1} \alpha_s^{(k+1)} t^s = \sum_{s=0}^{k} \alpha_s^{(k)} t^{s+1} - a_k \sum_{s=0}^{k} \alpha_s^{(k)} t^s - b_{k-1} \sum_{s=0}^{k-1} \alpha_s^{(k+1)} t^s .$$

Hence, upon equating coefficients of like powers we arrive at recurrence relations for the coefficients of $\bar{P}_{k+1}(t)$:

$$\alpha_{k+1}^{(k+1)} = \alpha_k^{(k)} = 1$$

$$\alpha_k^{(k-1)} = 0$$

$$\alpha_s^{(k+1)} = \alpha_{s-1}^{(k)} - a_k\alpha_s^{(k)} - b_{k-1}\,\alpha_s^{(k-1)} \qquad (5.19)$$

$$s = 0,1,\ldots,k.$$

These formulas express the coefficients of $\bar{P}_{k+1}(t)$ in terms of the scalar products , $c_i = (Z_{i-k}, Z_k)$.

Let ξ_λ denote, as before, the spectral function of A, so that

$$A = \int_m^M \lambda d\xi_\lambda$$

with m and M the glb and lub of the spectrum of A, respectively. Then

$$p_k = \bar{P}_k(A)Z_0 = \int_m^M \bar{P}_k(\lambda)d\xi_\lambda Z_0,$$

and, since the elements, p_k, are mutually orthogonal, we have

$$(p_i,p_k) = \int_m^M \bar{P}_i(\lambda)\bar{P}_k(\lambda)d(\xi_\lambda Z_0,Z_0) = 0 \qquad (5.20)$$
$$i \neq k$$

Thus the polynomials $\bar{P}_n(t)$ are orthogonal on the interval $[m,M]$ with respect to the non-decreasing function $(\xi_\lambda Z_0, Z_0)$ in the sense of (5.20). Hence, it can be shown that

(1) the polynomials $\bar{P}_n(t)$ form a Sturm series. That is, their roots, which are the eigenvalues of A_n, are real, distinct,

and separate one another,

(2) the roots all lie in the interval $[m,M]$,

(3) on any portion of the range of λ where $(\xi_\lambda Z_0, Z_0)$ is a constant, $\bar{P}_n(\lambda)$ can have no more than one root for each value of n.

It can be shown that when the method of moments is used to approximate the spectrum of a self-adjoint operator, as in this section, the eigenvalues which are largest in absolute value are determined most accurately. Suppose we wish to fine the eigenvalues which lie in a certain interval $[a,b]$. This can be done with the help of the next theorem.

Theorem 5.3 If A is an operator and P is a polynomial, then

$$\sigma(P(A)) = P(\sigma(A)) = \{P(\lambda) : \lambda \in \sigma(A)\}.$$

Proof:

For any complex number λ_0 there exists a polynomial Q such that $P(\lambda) - P(\lambda_0) = (\lambda - \lambda_0)Q(\lambda)$. It follows that $P(A) - P(\lambda_0 I) = (A - \lambda_0 I)\, Q(A)$. Thus, if $\lambda_0 \in \sigma(A)$ then $B = (A - \lambda_0 I)Q(A)$ is not invertible. To see this, suppose B is invertible. Then

$$(A - \lambda_0 I)Q(A)B^{-1} = BB^{-1} = I = B^{-1}B$$
$$= B^{-1}(A - \lambda_0 I)Q(A) = B^{-1}Q(A)(A - \lambda_0 I).$$

Therefore if B is invertible then $(A - \lambda_0 I)$ is invertible, which is a contradiction. Thus $P(A) - P(\lambda_0 I)$ is not invertible and we have proved that $P(\lambda_0) \in \sigma(P(A))$ and hence $P(\sigma(A)) \subseteq \sigma(P(A))$.

Now suppose $\lambda_0 \in \sigma(P(A))$ and let $\lambda_1, \lambda_2, \ldots \lambda_n$ be the roots of the equation

$$P(\lambda) = \lambda_0.$$

Thus,

$$P(A) - \lambda_0 I = \alpha(A - \lambda_1 I) \;\ldots\ldots (A - \lambda_n I) = 0$$

for some non-zero complex number α. Thus $(A-\lambda_j I)$ must fail to be invertible for at least one number j, $1 \le j \le n$. For such a value of j we have $\lambda_j \in \sigma(A)$ and $P(\lambda_j) = \lambda_0$, so that $\lambda_0 \in P(\sigma(A))$. Therefore $\sigma(P(A)) \subseteq P(\sigma(A))$ and thus $\sigma(P(A)) = P(\sigma(A))$. This completes the proof of the theorem.

Thus, to determine the eigenvalues of A in [a,b] we find a polynomial S(t) such that S(a+b/2) = 1, S(t) assumes its maximum value in [a,b], and S(t) deviates from zero by as little as possible outside [a,b]. Then, in view of Theorem (5.3), it is clear that the numerically largest eigenvalues of S(t) correspond exactly to those eigenvalues of A in [a,b]. In other words, if γ is an eigenvalue of S(A) then the number, or numbers, α such that $S(\alpha) = \gamma$ and $\alpha \in [a,b]$ are the desired eigenvalues of A.

D. Problems of the First Kind.

In this section, the solution of the problem of moments will be used to construct a sequence of operators converging to an inverse operator A^{-1}. Throughout this section it will be assumed that A^{-1} is bounded, thus there exists an interval $[-\sigma,\sigma]$ containing no part of the spectrum of A.

Consider a problem of the first kind,

$$Ax = f, \tag{5.21}$$

with A a bounded, self-adjoint operator and f an element of H_Z. To solve this problem, select a function, $\varphi(t)$, continuous on the interval $|t| \leq ||A||$ and equal to t^{-1} outside $[-\sigma,\sigma]$. Since the spectrum of A falls outside $[-\sigma,\sigma]$ it follows that $\varphi(A) = A^{-1}$. Now, since $A_n \to A$ in H_Z it follows that

$$\varphi(A_n) \to A^{-1}$$

in H_Z (see Theorem 5.2). Therefore, the sequence $\{x_n\}$, where

$$x_n = \int_m^M \varphi(\lambda) d\xi_\lambda^{(n)} f \tag{5.22}$$

converges strongly to the solution of (5.21).

Let $\lambda_k^{(n)}$ and $x_k^{(n)}$ denote the eigenvalues and eigenfunctions of A_n. Then

$$\xi_\lambda^{(n)} Z = \sum_k (Z, x_k^{(n)})\, x_k^{(n)},$$

where the summation extends over all k for which $\lambda_k < \lambda$. Thus, formula (5.22) becomes

$$x_n = \sum_{k=0}^{n-1} \varphi(\lambda_k^{(n)})(f, x_k^{(n)}) x_k^{(n)}. \tag{5.23}$$

Equation (5.23) can be used to obtain the approximate solution, x_n, regardless of the spectral distribution. To obtain x_n it is necessary to determine the spectrum of A_n. This can be avoided if the spectral distribution of A is known.

Consider the case where A is a positive definite self-adjoint operator, and let its spectrum lie in the interval $[m,M]$, with $m > 0$. Then the eigenvalues of A_n are all positive and lie in the interval $[m,M]$, in which $\varphi(t) = t^{-1}$. Therefore

$$x_n = \int_m^M \varphi(\lambda) d\xi_\lambda^{(n)} f = \int_m^M 1/\lambda \; d\xi_\lambda^{(n)} f.$$

Now A_n is an operator on the subspace H_n and $\xi_\lambda^{(n)}$ projects each element of H_Z into H_n. Thus, if f_n denotes the projection of f on H_n, then.,

$$\xi_\lambda^{(n)} f = \xi_\lambda^{(n)} f_n$$

and

$$x_n = \int_m^M 1/\lambda \; d\xi_\lambda^{(n)} f_n = A_n^{-1} f_n.$$

Thus, x_n satisfies the equation

$$A_n x_n = f_n \tag{5.24}$$

whose solution according to Chapter III is given by

$$x_n = A_n^{-1}[\, F_{n-1}(A_n) - \frac{F_{n-1}(0)}{P_n(0)} P_n(A_n)]Z_0,$$

with

$F_{n-1}(A_n)Z_0 = f_n$. But x_n is a polynomial of degree (n-1) in A_n:

$$x_n = Q_{n-1}(A_n)Z_0,$$

where

$$Q_{n-1}(t) = 1/t\,[F_{n-1}(t) - \frac{F_{n-1}(0)}{P_n(0)} P_n(t)].$$

Thus, from the way in which A_n is constructed,

$$Q_{n-1}(A_n)Z_0 = Q_{n-1}(A)Z_0$$

and

$$x_n = A^{-1}[\, F_{n-1}(A) - \frac{F_{n-1}(0)}{P_n(0)} P_n(A)]Z_0. \tag{5.25}$$

It can be shown that the approximation, x_n, can be determined in the case of a positive definite self-adjoint operator by the following algorithm:

$$x_{n+1} = x_n + h_n F_n(0) g_n \tag{5.26}$$

where

$$g_n = r_n + l_{n-1} g_{n-1} \qquad l_{n-1} = \frac{||r_n||^2}{||r_{n-1}||^2}$$

$$r_n = r_{n-1} - h_{n-1} A g_{n-1}$$

$$h_{n-1} = \frac{||r_{n-1}||^2}{(Ag_{n-1}, g_{n-1})} \qquad F_n(0) = F_{n-1}(0) + \frac{(f, r_n)}{||r_n||^2}$$

$$r_0 = g_0 = Z_0 \qquad F_0(0) = \frac{(f, Z_0)}{||Z_0||^2}.$$

We summarize the above results as

Theorem 5.4 Let A be a bounded self-adjoint positive(negative) definite operator on the Hilbert space H, and let Z_0 be any element of H. Then a sequence $\{x_n\}$ exists, obtainable by (5.25), or, equivalently, (5.26), which converges strongly to the solution of (5.21), provided f is in the subspace H_Z.

If we let $Z_0 = f$, then (5.25) simplifies to

$$x_n = A^{-1}\left[I - \frac{P_n(A)}{P_n(0)}\right] f \tag{5.25'}$$

and the algorithm described by (5.26) becomes

$$x_{n+1} = x_n + h_n g_n \tag{5.26'}$$

where

$$g_n = r_n + l_{n-1} g_{n-1} \qquad l_{n-1} = \frac{||r_n||^2}{||r_{n-1}||^2}$$

$$r_n = r_{n-1} - h_{n-1} A g_{n-1} \qquad r_0 = g_0 = f$$

$$h_{n-1} = \frac{||r_{n-1}||^2}{(Ag_{n-1}, g_{n-1})} \qquad f - Ax_n = r_n.$$

The advantage that the more general formulas (5.25) and (5.26) have over (5.25') and (5.26') is that if an approximate solution has been found for one right-hand side, f, then the solution for any other f can be determined directly without having to recompute the polynomials $P_n(A)$, g_n and r_n.

Suppose that A is a bounded self-adjoint operator having a spectrum with the property that on the left side of the origin is consists of a finite number of eigenvalues which we let be

$$\lambda_1 < \lambda_2 \leq \ldots < \lambda_k < -\sigma.$$

Concerning the right side of the origin, nothing further is assumed.

Since none of the spectrum of A lies in $[-\sigma,\sigma]$, the function $(\xi_\lambda Z_0, Z_0)$ is constant in this interval. Thus, for each n, no more than one root of $P_n(t)$ can lie in $[-\sigma,\sigma]$, and with increasing n this root will tend to $\lambda_k < -\sigma$. Hence, from some value of n onward, this root will lie outside $[-\sigma,\sigma]$. Thus, for such values of n, the approximate solution, x_n, of equation (5.21) will be given by (5.22):

$$x_n = \int_m^M \varphi(\lambda) d\xi_\lambda^{(n)} f = \int_m^M 1/\lambda \; d\xi_\lambda^{(n)} f = A_n^{-1} f_n .$$

Therefore x_n also satisfies

$$A_n x_n = f_n$$

and can be computed as before by means of (5.25) and (5.25').

Thus we have proved the following:

<u>Theorem 5.5</u> Let A be a bounded self-adjoint operator defined on the Hilbert space H and let A have a bounded inverse. If on one side of the origin, the spectrum of A consists of only a finite number of eigenvalues and on the other side it is bounded, then beginning with some value of n, the x_n defined by (5.25) and (5.25') exists and, with increasing n, the sequence $\{x_n\}$ converges strongly to the solution of equation (5.21) provided its right-hand side, f, is in the subspace H_Z.

Notice that the algorithms defined by (5.26) and (5.26') are not valid if A is not positive definite, because for certain values of n the approximate solution, x_n, need not exist, since for some n it could be that $P_n(0) = 0$, resulting in the coefficient h_n becoming infinite.

Now, suppose that A is a bounded self-adjoint operator having an infinite number of points in its spectrum to the left and to the right of the origin. Let

$$x_n = \varphi(A_n)f_n$$

where

$$f_n = \bar{f}_n + \bar{g}_n$$

with $\bar{f}_n$ an element of $\bar{H}_n$, the subspace generated by the eigenfunctions of A_n corresponding to eigenvalues lying outside $[-\sigma,\sigma]$, and $\bar{g}_n$ is an element of the subspace of eigenfunctions of A_n corresponding to its eigenvalues lying inside $[-\sigma,\sigma]$. The approximate solution, x_n, can be found as follows. Let δ_n be an eigenvalue of A_n lying in $[-\sigma,\sigma]$, and let

$$\bar{x}_n = A_n^{-1}[\ \bar{f}_n + C(A_n - \delta_n I)^{-1}P_n(A)Z_0] \tag{5.27}$$

Also let

$$\bar{f}_n = \bar{F}_{n-1}(A_n)Z_0$$

where $\bar{F}_{n-1}(t)$ is a polynomial of degree (n-1). Choose C so that

$$\bar{F}_{n-1}(0) - C\,\frac{P_n(0)}{\delta_n} = 0.$$

Thus, A_n may be replaced by A in (5.27) because it involves a polynomial of degree (n-1) in A_n, and by the way A_n is constructed

$A_n^k Z_0 = A^k Z_0$ for $k = 0,1,\ldots n-1$.

This leads to

$$\bar{x}_n = A^{-1}[\bar{F}_{n-1}(A) + \frac{\delta_n \bar{F}_{n-1}(0)}{P_n(0)}\,(A - \delta_n I)^{-1}P_n(A)]Z_0 \tag{5.28}$$

This formula need be used only if one of the roots of $\bar{P}_n(t)$ falls in $[-\sigma,\sigma]$.

F. Equations of the Second Kind

We now consider an equation of the second kind,

$$x = Ax + f \tag{5.29}$$

where A is a bounded self-adjoint operator, and $f \in H$. We also require that the operator $(I - A)$ is positive-definite, which will be true if $||A|| < 1$. Then the approximate solutions, x_k, can be obtained by methods entirely analogous to the last section. The resulting algorithm is;

$$x_{n+1} = x_n + F_{n-1}(1)h_n g_n \tag{5.30}$$

where

$$g_n = r_n + l_{n-1}g_{n-1} \qquad r_n = r_{n-1} - h_{n-1}(g_{n-1} - Ag_{n-1})$$

$$h_{n-1} = \frac{||r_{n-1}||^2}{(g_{n-1} - Ag_{n-1}, g_{n-1})} \qquad l_{n-1} = \frac{||r_n||^2}{||r_{n-1}||^2}$$

$$F_n(1) = F_{n-1}(1) + \frac{(f, r_n)}{||r_n||^2} \qquad r_0 = g_0 = Z_0$$

$$F_0(1) = \frac{(f, Z_0)}{||Z_0||^2}\,.$$

Convergence has been proved only if A is a completely continuous self-adjoint operator, $(I - A)$ is positive definite, and $(I - A)^{-1}$ is bounded.

If $Z_0 = f$, then (5.30) becomes

$$x_{n+1} = x_n + h_n g_n \tag{5.30'}$$

where

$$g_n = r_n + l_{n-1}g_{n-1} \qquad r_n = r_{n-1} - h_{n-1}(g_{n-1} - Ag_{n-1})$$

$$h_{n-1} = \frac{||r_{n-1}||^2}{(g_{n-1} - Ag_{n-1}, g_{n-1})} \qquad l_{n-1} = \frac{||r_n||^2}{||r_{n-1}||^2}$$

$$r_0 = g_0 = f \qquad\qquad r_n = f - x_n + Ax_n.$$

However, if I - A is not positive definite but satisfies the conditions of Theorem 5.4, an approximate solution to equation (5.29) can be obtained in the form

$$x_n' = \sum_{k=0}^{n-1} \beta_k \rho_k;$$

where the β_k are determined from the system of equations:

$$-\beta_{k-1} + (1 - a_k)\beta_k - b_k\beta_{k+1} = \frac{(f,\rho_k)}{||\rho_k||^2}, \qquad (5.31)$$

$$\beta_{-1} = \beta_n = 0.$$

CHAPTER VI

OTHER APPLICATIONS

In this section we will show how the method of moments can be used to solve various time-dependent problems, and we will consider the difficulties involved when an attempt is made to apply the method of moments to problems involving unbounded operators.

A. Time-Dependent Problems

Theorem 6.1 Suppose A is a symmetric positive definite linear operator defined on a dense subset of H, and that A^{-1} is defined on H. Then the problem

$$\frac{\partial x}{\partial t} = - Ax, \quad x(0) = x_0 \in H, \tag{6.1}$$

can be solved by finding the sequence of solutions of

$$A_n \frac{\partial x_n}{\partial t} + x_n = 0, \; x_n(0) = x_0 \tag{6.2}$$

where A_n is the solution to a certain moment problem. The sequence $\{x_n\}$ converges strongly to x.

Proof:

We write equation (6.1) in the form

$$A^{-1} \frac{\partial x}{\partial t} + x = 0 \tag{6.1'}$$

Setting

$$\begin{aligned} Z_0 &= x_0 \\ Z_1 &= A^{-1} Z_0 \\ &\ldots\ldots\ldots\ldots \\ Z_n &= A^{-1} Z_{n-1} \\ &\ldots\ldots\ldots\ldots \end{aligned}$$

we construct the sequence $\{A_m\}$ of solutions of problems of moments. Notice that is not necessary to use A^{-1} to find Z_{k+1} from Z_k; instead we could solve the corresponding problems of the first kind, namely,

$$A\ Z_{k+1} = Z_k.$$

Now, we replace (6.1') by the equation

$$A_n \frac{\partial x_n}{\partial t} + x_n = 0, \quad x_n(0) = x_0, \tag{6.3}$$

and show that $x_n \to x$.

If we let ξ_λ denote the spectral function of the inverse operator the required solution is given by

$$x = \int_0^M e^{-t/\lambda}\ d\xi_\lambda x_0,$$

with M the lub of the spectrum of A^{-1}, and thus $M = ||A^{-1}||$.

To see this, note that

$$A^{-1}\left[\ \frac{\partial}{\partial t}\left(\int_0^M e^{-t/\lambda}\ d\xi_\lambda x_0\right)\right] + \int_0^M e^{-t/\lambda}\ d\xi_\lambda x_0$$

$$= A^{-1}\left[\int_0^M - 1/\lambda\ e^{-t/\lambda}\ d\xi_\lambda x_0\right] + \int_0^M e^{-t/\lambda}\ d\xi_\lambda x_0$$

$$= \int_0^M \lambda d\xi_\lambda\left[\int_0^M -1/\lambda\ e^{-t/\lambda}\ d\xi_\lambda x_0\right] + \int_0^M e^{-t/\lambda} d\xi_\lambda x_0$$

$$= \int_0^M \lambda(-1/\lambda)\ e^{-t/\lambda}\ d\xi_\lambda x_0 + \int_0^M e^{-t/\lambda}\ d\xi_\lambda x_0 = 0.$$

It is clear that this solution satisfies the initial condition

$$x(0) = x_0$$

since

$$x(0) = \int_0^M 1 \cdot\ d\xi_\lambda x_0 = x_0.$$

Since $e^{-t/\lambda}$ is continuous for $\lambda > 0$ for all $t \geq 0$ and since the strong convergence of a sequence of operators implies the strong convergence of a continuous function of them, it follows that the sequence $\{x_n\}$ of solutions of (6.3) converges strongly to x.

This completes the proof of the theorem.

We will now show how the approximate solution, x_n, can be computed. Set

$$x_n = N_0(t)Z_0 + N_1(t)Z_1 + \dots + N_{n-1}(t)Z_{n-1},$$

and substitute this in (6.3). Using (3.1) we obtain

$$N_0(t)Z_0 + (N_1 + \frac{dN_0}{dt})Z_1 + \dots + (N_{n-1} + \frac{dN_{n-2}}{dt})Z_{n-1}$$

$$+ \frac{dN_{n-1}}{dt} E_n Z_n = 0.$$

Now, substituting the expression given in (3.3) for $E_n Z_n$, we obtain

$$(N_0 - \alpha_0 \frac{dN_{n-1}}{dt}) Z_0 + (N_1 + \frac{dN_0}{dt} - \alpha_1 \frac{dN_{n-1}}{dt})Z_1 + \dots$$

$$\dots + (N_{n-1} + \frac{dN_{n-2}}{dt} - \alpha_{n-1} \frac{dN_{n-1}}{dt})Z_{n-1} = 0.$$

The elements, Z_k are, by assumption, linearly independent; otherwise, as indicated in the last paragraph of Chapter III, the approximation procedure would terminate earlier and yield the exact solution. Therefore, we can equate their coefficients to zero, yielding

$$N_0 - \alpha_0 \frac{dN_{n-1}}{dt} = 0$$

$$\frac{dN_0}{dt} + N_1 - \alpha_1 \frac{dN_{n-1}}{dt} = 0$$

$$\dots\dots\dots\dots\dots\dots\dots$$

$$\frac{dN_{n-2}}{dt} + N_{n-1} - \alpha_{n-1} \frac{dN_{n-1}}{dt} = 0, \tag{6.4}$$

with corresponding initial conditions

$$N_0(0) = 1,\ N_1(0) = N_2(0) = \dots = N_{n-1}(0) = 0.$$

Equations (6.4) may be solved by a classical method or by means of Laplace transforms. Let

$$\xi_k = \int_0^\infty N_k(t)\, e^{-\lambda t}\, dt\,.$$

Then

$$\xi_0 - \alpha_0\lambda\xi_{n-1} = 0$$

$$\lambda\xi_0 + \xi_1 - \alpha_1\lambda\xi_{n-1} = 0$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\lambda\xi_{n-2} + \xi_{n-1} - \alpha_{n-1}\lambda\xi_{n-1} = 0,$$

and hence

$$\xi_j = \frac{P_{n-j-1}(-1/\lambda)}{\lambda^2 P_n(-1/\lambda)}\,, \; j \geq 1,$$

$$\xi_0 = \frac{\alpha_0}{\lambda P_n(-1/\lambda)}\,,$$

where $P_n(\lambda)$ is as in (3.5).

The solution to (6.4) is thus given by

$$N_j(t) = 1/2\pi i \int_{\sigma - i\infty}^{\sigma + i\infty} \frac{P_{n-j-1}(-1/\lambda)}{\lambda^2 P_n(-1/\lambda)} e^{\lambda t} d\lambda, \quad j \geq 1$$

$$N_0(t) = 1/2\pi i \int_{\sigma - i\infty}^{\sigma + i\infty} \frac{\alpha_0 e^{\lambda t}}{\lambda P_n(-1/\lambda)}\, d_\lambda, \tag{6.5}$$

where σ is any number not in the spectrum of A.

Therefore, to construct x_n it is necessary to solve the equations

$$AZ_{k+1} = Z_k,\; k = 0, 1, \ldots, n-1$$

then use the Z_k thus determined to solve system (3.7) repeatedly to find the polynimials $P_k(\lambda)$, $k = 0, 1, \ldots, n$, and then use these in (6.5) to determine the coefficients.

<u>Theorem 6.2</u> Suppose A satisfies the conditions of Theorem 6.1 and

$$\frac{\partial x}{\partial t} = - Ax + gf(t),\ x(0) = 0, \tag{6.6}$$

with $g \in H$ and $f\ (t)$ a scalar function. Then the sequence $\{x_n\}$ of solutions of

$$A_n \frac{\partial x_n}{\partial t} + x_n = A^{-1}\, gf(t), \tag{6.7}$$

with A_n the solution to the problem of moments defined by

$$Z_0 = g,$$
$$Z_1 = A^{-1}g,$$
$$\cdots\cdots\cdots$$
$$Z_n = A^{-1}\, Z_{n-1},\ \ldots,$$

converges strongly to x.

Proof:

The solution to (6.6) is

$$x = \int_0^M \left[\int_0^t f(w)e^{-(t\ -\ w)/\lambda}\, dw\right] d\xi_\lambda g \tag{6.8}$$

and since $e^{-(t\ -\ w)/\lambda}$ is continuous for $\lambda > 0$ the same argument as used in Theorem 6.1 verifies that the sequence $\{x_n\}$ converges strongly to x. This completes the proof of the theorem.

To see that x is actually the solution note that $x(0) = 0$ and

$$A^{-1}\frac{\partial x}{\partial t} = \int_0^M \lambda d\xi_\lambda \frac{\partial}{\partial t}\left\{\int_0^M \left[\int_0^t f(w)\ e^{-(t-w)/\lambda} dw\right] d\xi_\lambda g\right\}$$

$$= \int_0^M \lambda d\xi_\lambda \left\{\int_0^M \left[\frac{\partial}{\partial t}\int_0^t f(w)e^{-(t-w)/\lambda} dw\ \right] d\xi_\lambda g\right\}$$

$$= \int_0^M \lambda d\xi_x \{ \int_0^M [\int_0^t -1/\lambda \ f(w) e^{-(t-w)/\lambda} dw + f(t)] d\xi_\lambda g \}$$

$$= \int_0^M \lambda [\int_0^t -1/\lambda \ f(w) e^{-(t-w)/\lambda} dw + f(t)] d\xi_\lambda g$$

$$= \int_0^M [\int_0^t -f(w) e^{-(t-w)/\lambda} dw + \lambda f(t)] d\xi_\lambda g$$

$$= - \int_0^M [\int_0^t f(w) e^{-(t-w)/\lambda} dw] d\xi_\lambda g + \int_0^M \lambda f(t) d\xi_\lambda g$$

$$= - x + \int_0^M \lambda d\xi_\lambda [g f(t)]$$

$$= - x + A^{-1} g f(t)$$

Therefore

$$A^{-1} \frac{\partial x}{\partial t} + x = A^{-1} g f(t).$$

We now show how the approximate solution, x_n, can be found. Let

$$x_n = N_0(t) Z_0 + N_1(t) Z_1, + \dots + N_{n-1}(t) Z_{n-1}.$$

The substitution of this in equation (6.7) yields,

$$N_0 - \alpha_0 \frac{dN_{n-1}}{dt} = 0$$

$$\frac{dN_0}{dt} + N_1 - \alpha_1 \frac{dN_{n-1}}{dt} = f(t)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{dN_{n-2}}{dt} + N_{n-1} - \alpha_{n-1} \frac{dN_{n-1}}{dt} = 0,$$

which can be solved as above by Laplace transforms, yielding

$$N_j = 1/2\pi i \int_{\sigma - i\infty}^{\sigma + i\infty} \frac{F(\lambda) \ P_{n-j-1}(-1/\lambda)}{\lambda^2 P_n(-1/\lambda)} \ e^{\lambda t} d\lambda, \ j \geq 1$$

$$N_0 = 1/2\pi i \int_{\sigma - i\infty}^{\sigma + i\infty} \frac{\alpha_0 e^{\lambda t} F(\lambda)}{\lambda P_n(-1/\lambda)} \ d\lambda,$$

where

$$F(\lambda) = \int_0^\infty f(t)\, e^{-\lambda t} dt.$$

Notice that constructing the approximate solution, x_n, involves calculating the integral $F(\lambda)$ and determining the polynomials $P_n(\lambda)$, as above, to be used in calculating the elements, N_j.
As an example of the above consider the temperature distribution in a rod. The temperature distribution satisfies the equation

$$\frac{\partial \phi}{\partial t} = a(x) \frac{\partial^2 \phi}{\partial x^2}, \tag{6.9}$$

and we will consider the case where

$$a(x) = 1/(1+x^2),$$

the initial temperature is

$$\phi(x,0) = (1 - x^2)/2,$$

and at the ends of the rod

$$\phi(-1,t) = \phi(1,t) = 0.$$

Thus, if we let

$$A\phi = - \frac{1}{1 + x^2} \frac{\partial^2 \phi}{\partial x^2}$$

then we can write (6.9) as

$$\frac{\partial \phi}{\partial t} = - A\phi. \tag{6.9'}$$

It can be shown that A is positive definite. The sequence of functions, $\{Z_k\}$ is calculated from the formula

$$Z_0(x) \equiv \phi(x,0) = (1-x^2)/2,$$

$$Z_k(x) = A^{-1} Z_{k-1}$$

or

$$Z_k''(x) = - (1 + x^2) Z_{k-1}(x), \quad Z_k(-1) = Z_k(1) = 0.$$

Integrating the equations, we obtain

$$Z_0(x) = 1/2(1 - x^2)$$

$$Z_1(x) = 1/60(x^6 - 15x^2 + 14)$$

$$Z_2(x) = -1/3360(x^8 - 70\ x^4 + 392\ x^2 - 323)$$

....

We define the scalar product

$$[Z_i,\ Z_k] = \int_{-1}^{1} Z_i''(x)\ Z_k(x)\ dx.$$

Applying the formulas (6.4) and (6.5) we can write down the equation leading to the first approximation:

$$n = 1:\ \phi,(x,t) = N_0(t)Z_0(x),$$

$$N_0 - \alpha_0 \frac{dN_0}{dt} = 0,\ N_0(0) = 1,$$

$$\alpha_0 = \frac{-[Z_1, Z_0]}{[Z_0,\ Z_0]} = -\ 0.4571,$$

and those for the second approximation:

$$n = 2:\quad \phi_2(x,t) = N_0(t)Z_0(x) + N_1(t)Z_1(x),$$

$$N_0 - \alpha_0 \frac{dN_1}{dt} = 0$$

$$\frac{dN_0}{dt} + N_1 - \alpha_1\ \frac{dN_1}{dt} = 0$$

$$N_0(0) = 1,\ N_1(0) = 0,$$

$$\alpha_0 = 10^{-2}(2.1781),\quad \alpha_1 = -\ 0.50674.$$

On solving the resulting differential equations, we find that

$$\phi_1(x,t) = Z_0(x)\ e^{-2.187t}$$

and

$$\phi_2(x,t) = [2.428Z_1(x) - 0.1151Z_0(x)]\ e^{-2.177t}$$
$$- [2.428Z_1(x) - 1.1151Z_0(x)]e^{-2.109}.$$

B. Unbounded Operators

Suppose A is an unbounded symmetric linear operator defined on a linear manifold, L_A, dense in H. To formulate the problem of moments, we first must find an element $Z_0 \in H$ on which all powers of A are defined. There may not be such an element, but we will suppose there is one. By solving the moments problem, we can construct a sequence $\{A_n\}$ of operators such that

$$Z_k = A^k Z_0 = A_n^k Z_0, \quad k = 0, 1, \ldots, n-1$$

$$E_n Z_n = A_n^n Z_0.$$

The linear manifold $L_{Z_{0(m)}}$ (Definition 35) is clearly a subset of the domain of H. Furthermore, the sequence $\{A_n\}$ converges strongly to A on $L_{Z_{0(m)}}$, since for $n \geq m+1$

$$Ax = A_n x,$$

for $x \in L_{Z_{0(m)}}$.

Thus,

$$A' = \lim_{n \to \infty} A_n$$

coincides with A on the linear manifold $L_{Z_{0(m)}}$.

If $\overline{A}'$, the closure of A', is self-adjoint as is the closure of A, then it can be shown that the sequence of spectral functions $\{\xi_\lambda^{(n)}\}$, corresponding to the operators A_n converges strongly to the spectral function of A.

In this case, the problem of moments is said to be determinate.

Thus, in attempting to apply the method of moments in this case we may be stopped immediately if no element Z_0 exists on which all powers of A are defined. Then, even if such an element exists, the problem of moments may not be determinate.

We now consider one case involving an unbounded operator which can be solved by the method of moments by reducing it to a linear problem with a bounded operator.

<u>Theorem 6.3</u> Suppose A is an unbounded linear operator defined on a linear manifold, L_A, which is dense in H, and

$$A = A_0 - K,$$

where A_0 is a positive definite self-adjoint operator defined on L_A and K is such that

$$T = A_0^{-1}K \tag{6.10}$$

is completely continuous in H_0, or can be extended to be completely continuous in H_0. Then the equation

$$Ax = f \tag{6.11}$$

can be converted to a linear equation involving the bounded operator T and can thus be solved by the method of moments.

Proof:

The linear manifold L_A is dense in H_0 and all elements of H_0 also belong to H. Thus equation (6.11) can be replaced by its equivalent

$$x - Tx = A_0^{-1} f, \tag{6.12}$$

which is already solvable by the method of moments [see (4.1)].

Next we will work one example problem using the method of moments.

Consider the problem

$$y = \int_0^1 K(x,\xi)y(\xi)d\xi + x, \tag{6.13}$$

where

$$K(x,\xi) = \begin{cases} \xi(1-x), & \xi \le x \\ x(1-\xi), & \xi \ge x. \end{cases}$$

Let

$$Ay = \int_0^1 K(x,\xi)y(\xi)d\xi.$$

Then the problem can be written as

$$y = Ay + x.$$

The operator A is completely continuous and self-adjoint, therefore the method given in Chapter IV, Section B, may be used in solving (6.13).

Let

$$Z_0 = x$$

$$Z_1 = AZ_0 = 1/6\ x\ (1-x^2)$$

$$Z_2 = AZ_1 = 1/360\ x\ (7-10x^2 + 3x^4)$$

$$Z_3 = 1/15120\ x\ (31 - 49x^2 + 21x^4 - 3x^6)$$

The inner product is defined by

$$(Z_i, Z_j) = \int_0^1 Z_i(x)Z_j(x)dx.$$

The resulting inner products are:

$$(Z_0, Z_0) = .33333333$$

$$(Z_0, Z_1) = .02222222 = (Z_1, Z_0)$$

$$(Z_0, Z_2) = .00211640 = (Z_2, Z_0) = (Z_1, Z_1)$$

$$(Z_0, Z_3) = .00021164 = (Z_3, Z_0) = (Z_1, Z_2) = (Z_2, Z_1)$$

$$(Z_1, Z_3) = .00002137 = (Z_3, Z_1) = (Z_2, Z_2)$$

$$(Z_2, Z_3) = .00000216$$

Then these inner products are used in system (3.7) to determine the numbers α_0, α_1, α_2. They are

$$\alpha_0 = .00002038$$

$$\alpha_1 = - .00019413$$

$$\alpha_2 = -.10117292$$

Thus the number $P_3(1)$ as defined by (3.5) becomes

$$P_3(1) = 1 + \alpha_0 + \alpha_1 + \alpha_2 = .89865333$$

The coefficients a_0, a_1, a_2 defined by (3.17) are

$$a_0 = .99997733$$

$$a_1 = 1.00019335$$

$$a_2 = 1.11277616$$

Therefore the third approximation as defined by (3.17) becomes

$$\begin{aligned} x_3 &= a_0 Z_0 + a_1 Z_1 + a_2 Z_2 \\ &= .00927313x^5 - .19760932x^3 + 1.18831351x. \end{aligned} \tag{6.14}$$

The following table indicates how this solution compares with the exact solution, which is sin x/sin 1.

x	x_3	Exact Solution
0	0	0
.1	.1190	.1186
.2	.2361	.2360
.3	.3512	.3512
.4	.4628	.4628
.5	.5697	.5697
.6	.6710	.6710
.7	.7656	.7656
.8	.8525	.8525
.9	.9309	.9309
1.0	1.0000	1.0000

BIBLIOGRAPHY

1. Akhiezer, N. I. and Glazman, I. M., Theory of Linear Operators In Hilbert Space, vol I, Frederick Ungar Publishing Co., New York, 1961.

2. Akhiezer, N. I. and Glazman, I. M., Theory of Linear Operators In Hilbert Space, vol II, Frederick Ungar Publishing Co., New York, 1963.

3. Danese, A. E., Advanced Calculus, vol I, Allyn and Bacon, Inc., Boston, 1965.

4. Dunford, N. and Schwartz, J.T., Linear Operators, Part II: Spectral Theory in Hilbert Space, Interscience Publications, New York, 1963.

5. Halmos, P., Finite Dimensional Vector Spaces, Princeton University Press, London, 1942.

6. Halmos, P., Introduction to Hilbert Space and the Theory of Spectral Multiplicity, Chelsea Publishing Co., New York, 1951.

7. Lorch, E. R., Spectral Theory, Oxford University Press, New York, 1962.

8. Mikhlin, S. G., Variational Methods in Mathematical Physics, The MacMillan Co., New York, 1964.

9. Rudin, W. Principles of Mathematical Analysis, McGraw-Hill Book Co., New York, 1964.

10. Stone, M. H., Linear Transformations in Hilbert Space and Their Applications To Analysis, American Mathematical Society, New York, 1964.

11. Vorobyev, Y. V., Method of Moments in Applied Mathematics, Gordon and Breach Science Publishers, New York, 1965.